AD NO. ___________________________________
ATEC PROJECT NO. 2011-DT-ATC-DODSP-F0292
REPORT NO. ATC 11592

STANDARDIZED

UXO TECHNOLOGY DEMONSTRATION SITE

SCORING RECORD NO. 944

SITE LOCATION:
ABERDEEN PROVING GROUND

DEMONSTRATOR:
WHITE RIVER TECHNOLOGIES
115 ETNA ROAD, BUILDING 3, SUITE 1
LEBANON, NH 03766

TECHNOLOGY TYPE/PLATFORM:
OPTEMA
TOWED ARRAY

AREAS COVERED:
BLIND GRID
OPEN FIELD (INDIRECT FIRE)

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

AUGUST 2014

Prepared for:
SERDP/ESTCP
MUNITIONS MANAGEMENT
ALEXANDRIA, VA 22350

U.S. ARMY TEST AND EVALUATION COMMAND
ABERDEEN PROVING GROUND, MD 21005-5001

DISTRIBUTION UNLIMITED, AUGUST 2014.

DEPARTMENT OF THE ARMY
U.S. ARMY ABERDEEN TEST CENTER
400 COLLERAN ROAD
ABERDEEN PROVING GROUND, MARYLAND 21005-5059

REPLY TO
ATTENTION OF

TEDT-AT-SL-M

MEMORANDUM FOR Program Manager – SERDP/ESTCP, Munitions Management, Mr. Herb Nelson, 4800 Mark Center Drive, Suite 17D08, Alexandria, VA 22350-3600

SUBJECT: Standardized Unexploded Ordinance Technology (UXO) Demonstration Site Scoring Record No. 944

1. The subject Scoring Record is submitted for your information and retention.

2. The point of contact for this office is Mr. Leonard Lombardo, Test Officer, Survivability/Lethality Directorate, Maritime/Threat Detection Systems Survivability Division (TDSS) and may be reached at 410-278-1286 or via e-mail, leonard.c.lombardo.civ@mail.mil.

TRUTH IN TESTING!

FOR THE COMMANDER:

SHEPPARD.TRACY.V.1200534219

Digitally signed by SHEPPARD.TRACY.V.1200534219
DN: c=US, o=U.S. Government, ou=DoD, ou=PKI, ou=USA, cn=SHEPPARD.TRACY.V.1200534219
Date: 2014.11.09 20:03:28 -05'00'

Encl

TRACY V. SHEPPARD
Director, Survivability/Lethality Directorate

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
August 2014	Final	April 2014

4. TITLE AND SUBTITLE	
Standardized UXO Technology Demonstration Site Scoring Record No. 944	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER** 2011-DT-ATC-DODSP-F0292
Hill, Scott Lombardo, Leonard Fling Jr., Rick	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Commander U.S. Army Aberdeen Test Center ATTN: TEDT-AT-SLE Aberdeen Proving Ground, MD 21005-5059	ATC-11592
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**
Program Manager SERDP/ESTCP Munitions Management Alexandria, VA 22350	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)** Same as Item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution unlimited.

13. SUPPLEMENTARY NOTES

None

14. ABSTRACT

This scoring record documents the efforts of the Environmental Security Technology Certification Program to detect and discriminate inert unexploded ordnance (UXO) utilizing the APG Standardized UXO Technology Demonstration Site calibration lanes and open field sites. This Scoring Record was coordinated by J. Stephen McClung and the Standardized UXO Technology Demonstration Site Scoring Committee. Organizations on the committee include the U.S. Army Corps of Engineers, the Environmental Security Technology Certification Program, the Strategic Environmental Research and Development Program, the Institute for Defense Analysis, the U.S. Army Environmental Command, and the U.S. Army Aberdeen Test Center.

15. SUBJECT TERMS

Environmental Security Technology Certification Program, APG, Standardized UXO Technology Demonstration Site, blind grid, open field, indirect fire, towed array.

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
Unclassified	Unclassified	Unclassified	SAR		

ACKNOWLEDGMENTS

Authors:

Leonard Lombardo
Threat Detection and Systems Survivability (TDSS)
U.S. Army Aberdeen Test Center (ATC)
Aberdeen Proving Ground (APG)

Rick Fling
Aberdeen Test Support Services (ATSS)
Sverdrup Technology, Inc.
Aberdeen Proving Ground (APG)

Contributors:

Scott Hill
Threat Detection and Systems Survivability (TDSS)
U.S. Army Aberdeen Test Center
Aberdeen Proving Ground

TABLE OF CONTENTS

SECTION 4. TECHNICAL PERFORMANCE RESULTS

PAGE

SECTION 5. APPENDIXES

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of unexploded ordnance (UXO) require testing so that their performance can be characterized. To that end, Standardized Test Sites have been developed at Aberdeen Proving Ground (APG), Maryland, and U.S. Army Yuma Proving Ground (YPG), Arizona. These test sites provide a diversity of geology, climate, terrain, and weather as well as diversity in ordnance and clutter. Testing at these sites is independently administered and analyzed by the Government for the purposes of characterizing technologies, tracking performance with system development, comparing performance of different systems, and comparing performance in different environments (app E, ref 1).

The Standardized UXO Technology Demonstration Site Program is a multiagency program spearheaded and funded by the Environmental Securities Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP). The U.S. Army Aberdeen Test Center (ATC) provides programmatic and field support for technology demonstration and evaluation, and maintains a repository of inert munition items available to the UXO community. The U.S. Army Environmental Command maintains the Standardized UXO Technology Demonstration Site Program web page (http://aec.army.mil/usaec/technology/uxo01.html), which contains program information, vendor demonstration instructions and copies of all published vendor demonstration scoring records.

1.2 SCORING OBJECTIVES

The objective in the Standardized UXO Technology Demonstration Site Program is to evaluate the detection and discrimination capabilities of a given technology under various field and soil conditions. Inert munitions and clutter items are positioned in various orientations and depths in the ground.

The evaluation objectives are as follows:

a. To determine detection and discrimination effectiveness under realistic scenarios with various targets, geology, clutter, density, topography, and vegetation.

b. To determine cost, time, and workforce requirements to operate the technology.

c. To determine the demonstrator's ability to analyze survey data in a timely manner and provide prioritized Target Lists with associated confidence levels.

d. To provide independent site management to enable the collection of high quality, ground-truth (GT), geo-referenced data for post-demonstration analysis.

1.2.1 <u>Scoring Methodology</u>

a. The scoring of the demonstrator's performance is conducted in two stages: response stage and discrimination stage. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver-operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of clutter detection (P_{cd}) or the probability of false positive (P_{fp}). Those that do not correspond to any known item are termed background alarms. The background alarms are addressed as either probability of background alarm (P_{ba}) or background alarm rate (BAR).

b. The response stage scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate munitions from other anomaly sources. For the blind grid response stage, the demonstrator provides a target response from each and every grid square along with a threshold below which target responses are deemed insufficient to warrant further investigation. This list is generated with minimal processing and, since a value is provided for every grid square, includes amplitudes both above and below the system noise level. For the open field, the demonstrator provides a list of all anomalies deemed to exceed a demonstrator selected target detection threshold. An item (either munition or clutter) is counted as detected if a demonstrator indicates an anomaly within a specified distance (Halo Radius (R_{halo})) of a GT item.

c. The discrimination stage evaluates the demonstrator's ability to correctly identify munitions as such and to reject clutter. For the blind grid discrimination stage, the demonstrator provides the output of the discrimination stage processing for each grid square. For the open field, the demonstrator provides the output of the discrimination stage processing for anomaly reported in the response stage. The values in these lists are prioritized based on the demonstrator's determination that a location is likely to contain munitions. Thus, higher output values are indicative of higher confidence that a munitions item is present at the specified location. For digital signal processing, priority ranking is based on algorithm output. For other discrimination approaches, priority ranking may be based on rule sets or human judgment. The demonstrator also specifies the threshold in the prioritized ranking that provides optimum performance, (i.e., that is expected to retain all detected munitions and reject the maximum amount of clutter).

d. The demonstrator is also scored on efficiency and rejection ratios, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of munitions detections from the anomaly list, while rejecting the maximum number of anomalies arising from nonmunitions items. Efficiency measures the fraction of detected munitions retained after discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the maximum number of munitions detectable by the sensor and its accompanying clutter detection/false positive rate or BAR.

e. Based on configuration of the GT at the standardized sites and the defined scoring methodology, in some cases, there exists the possibility of having anomalies within overlapping halos and/or multiple anomalies within halos. In these cases, the following scoring logic is implemented:

(1) In situations where multiple anomalies exist within a single R_{halo}, the anomaly with the strongest response or highest ranking will be assigned to that particular GT item. If the responses or rankings are equal, then the anomaly closest to the GT item will be assigned to the GT item. Remaining anomalies are retained and scored until all matching is complete.

(2) Anomalies located within any R_{halo} that do not get associated with a particular GT item are excess alarms and will be disregarded.

f. In some cases, groups of closely spaced munitions have overlapping halos. The following scoring logic is implemented (app A, fig. A-1 through A-9):

(1) Overall site scores (i.e., P_d) will consider only isolated munitions and clutter items.

(2) GT items that have overlapping halos (both munitions and clutter) will form a group and groups may form chains.

(3) Groups will have a complex halos composed of the composite halos of all its GT items.

(4) Groups will have three scoring factors: groups found, groups identified, and group coverage. Scores will be based on 1:1 matches of anomalies and GT.

(a) Groups found (found). The number of groups that have one or more GT items matched divided by the total number of groups. Demonstrators will be credited with detecting a group if any item within the group is matched to an anomaly in their lists.

(b) Groups identified (ID). The number of groups that have two or more GT items matched divided by the total number of groups. Demonstrators will be credited with identifying that a group is present if multiple items within the composite halo are matched to anomalies in their lists.

(c) Group coverage (coverage). The number of GT items matched within groups divided by the total number of GT items within groups. This metric measures the demonstrator accuracy in determining the number of anomalies within a group. If five items are present and only two anomalies are matched, the demonstrator will score 0.4. If all five are matched, the demonstrator will score 1.0.

(5) Location error will not be reported for groups.

(6) Demonstrators will not be asked to call out groups in their scoring submissions. If multiple anomalies are indicated in a small area, the demonstrator will report all individual anomalies.

(7) Excess alarms within a halo will be disregarded.

g. All scoring factors are generated utilizing the Standardized UXO Probability and Plot Program, version 4.

1.2.2 Scoring Factors

Factors to be measured and evaluated as part of this demonstration include:

a. Response stage ROC curves:

(1) Probability of detection (P_d^{res}).

(2) Probability of clutter detection (P_{cd}).

(3) Background alarm rate (BAR^{res}) or probability of background alarm (P_{ba}^{res}).

b. Discrimination stage ROC curves:

(1) Probability of detection (P_d^{disc}).

(2) Probability of false positive (P_{fp}).

(3) Background alarm rate (BAR^{disc}) or probability of background alarm (P_{ba}^{disc}).

c. Metrics:

(1) Efficiency (E).

(2) False positive rejection rate (R_{fp}).

(3) Background alarm rejection rate (R_{ba}).

d. Other:

(1) Probability of detection by size, depth, and density.

(2) Classification by type (i.e., 20-, 40-, 105-mm, etc.).

(3) Location accuracy for single munitions.

(4) Equipment setup, calibration time, and corresponding worker-hour requirements.

(5) Survey time and corresponding worker-hour requirements.

(6) Reacquisition/resurvey time and worker-hour requirements (if any).

(7) Downtime due to system malfunctions and maintenance requirements.

SECTION 2. DEMONSTRATION

2.1 DEMONSTRATOR INFORMATION

2.1.1 Demonstrator Point of Contact (POC) and Address

POC: Jonathan Miller

Address: Etna Road, Building 3, Suite 1, Lebanon, NH 03766

2.1.2 System Description (Provided by Demonstrator)

The OPTEMA comprises an array of multi-directional transmitters and receivers that are optimally configured to provide good electromagnetic induction (EMI) characterization across the entire sensor swath. This capability is the basis for effective dynamic classification since sensor position during dynamic surveys is based on survey transects rather than on a priori target location. Consequently, it is likely that a large number of targets will be located at some lateral offset relative to the array center during a dynamic survey.

The OPTEMA sensor is built around the G&G Sciences National Instruments (NI)-based data acquisition framework. Data acquisition hardware is housed in a National Instruments PXI-1042 chassis and includes an NI PXI-8108 embedded controller (Windows OS) and six 8-channel 16-bit NI PXI-6143 A/D cards. Intermediate hardware is housed in an external module and includes the transmitter controller and power distribution board, and three 16-channel receiver boards. These components along with a 2000-W inverter are contained in a ruggedized vehicle-mounted chassis and compose the OPTEMA sensor electronics.

The OPTEMA sensor head comprises five transmitters and 14 three-axis receivers across a 1.8-m sensor swath. This design ensures that three orthogonal magnetic fields are produced at any across track location. The distribution of the 14 receivers also ensures that fields scattered by any target located across the sensor swath will be characterized sufficiently to constrain inversion of the data. The transmitter coils include four horizontal axis transmitters and one large vertical axis transmitter. The horizontal axis transmitters are wired in series pairs to provide two effective orthogonal excitation axes. The horizontal axis transmitters and receiver cubes share the same reference coordinate frame that is rotated 45^{o} from the principal coordinate system (i.e., referenced to the direction of travel).

Each OPTEMA transmitter produces an exponential current ramp that is rapidly terminated to generate a strong electromotive force in nearby targets. The receivers measure the decay of secondary magnetic field contributions over time gates spanning a period from 24 µs to up to approximately 8 ms. The combination of three effective transmitters (including horizontal axis transmitters in pairs) and 14 three-axis receivers provides 126 data channels. The magnetic field decays measured by the receivers usually include contributions from system components (e.g., supporting hardware, survey vehicle, etc.). To maximize sensitivity to anomalies, a background response is typically acquired to establish these intrinsic sources, and later subtracted from subsequent data sets to isolate the anomaly response. In addition to a background subtraction, a transmitter current normalization is applied to all data files to account for any small discrepancies in current between the three transmitters.

The OPTEMA uses a 50-percent duty cycle 60-Hz sub-harmonic as the base waveform. Data blocks (of duration BlockT) are built from a specified number of repeats (nRepeats) of this base waveform. The base waveform comprises a single period of a 50-percent duty cycle square wave voltage. During the voltage-on periods, the current in the energized transmitter coil ramps exponentially according to the coil time constant parameters. During the voltage-off periods, the receivers measure the ambient magnetic field decays. These decays are averaged with the appropriate sign changes for the positive and negative half cycles of the base period. Further signal averaging is applied for each repeat of the base waveform within the data block. In addition to the base repeats, the operator can select a number of blocks to stack (nStk) or average, in order to improve signal-to-noise ratio. Thus, the OPTEMA provides noise reduction through the inherent base repeat averaging and the explicit stack averaging.

Sensor head position and orientation data are provided by a Trimble RTK DGPS R8 receiver and a Microstrain 3DM-GX3-25 inertial sensor, respectively. Position and orientation sensors are mounted on the sensor tow-sled, which provides undercarriage protection and mobility for the sensor head. The sled can be towed by any mid-size utility vehicle (e.g., Kubota RTV-900 series) and can be mounted to the vehicle using a standard ball hitch. The tow-sled rides on a set of low-pressure pneumatic tires that provide a nominal 2.5-cm ground clearance for the sled skid plate. In uneven terrain, the skid plate will contact the ground occasionally.

Navigation for both static and dynamic measurements will be provided by a software navigation interface. Line segment files (for dynamic measurements) and cued files (for static measurements) will be generated prior to conducting surveys in the UXO Test Site areas. GT for the calibration grid will be used to generate cues for each of the items in the calibration grid. Grid coordinates for all three area objectives (calibration lanes, blind grid, and open-field indirect fire subarea) will be required to generate transect lines in each survey area. Format for each cue is a standard longitude, latitude coordinate in decimal degrees. Format for line segments is the longitude, latitude of each line endpoint in decimal degrees. A 1.2-m spacing for transect lines in all area objectives will be used. This provides a 33-percent overlap for adjacent transects. The target advance rate for all area objectives is 0.5 m/s (1.8 kph).

Figure 1. OPTEMA/TOWED ARRAY.

2.1.3 Data Processing Description (Provided by Demonstrator)

The section should be submitted for each area surveyed by the vendor. Discussion should include how target selection, parameter estimation, and classification vary by site area and objective. The following information should be submitted to ATC within 30 days before each area is surveyed:

a. Target Selection Criteria**:** This section will detail the target selection criteria and the data required to implement the criteria by answering the following questions:

(1) What kind of preprocessing (if any) is applied to the raw data (e.g., filtering, etc.)?

Background subtraction and current normalization steps are applied to the raw sensor data. Background response is determined in the calibration area and is subtracted from all subsequent data. For current normalization, the current in each transmitter coil is recorded in the data files. The raw data are then divided by the peak current in each coil to account for any variations in transmitter current over the course of the data collection period.

(2) What is the format of the data both pre- and post-processing of the raw data (e.g., ASCII, binary, etc.)?

Data are initially logged as .tem files and subsequently converted to .csv format for processing. After processing, the data are stored as .mat files.

(3) What algorithm is used for detection (e.g., peaks of signal surpassing threshold, etc.)?

A two-dimensional (2-D) map is created using the sum time decay for the Transmit-Z/Receive-Z data channels. Data are gridded and a 2-D interpolation is applied. A peak detection algorithm is applied to the map using a threshold based on the data noise floor standard deviation. A detection radius is applied to identify the region of interest (ROI) surrounding each peak. The radius size is based on the local gradient associated with the peak and the number of peaks associated with an anomaly (one peak for Z-data). If ROIs associated with multiple peaks overlap, a combined ROI is generated that encompasses the multiple detections. Finally, across track and along track indices are generated for each alarm in an ROI. These indices correspond to the receiver cube and sounding number associated with each alarm.

(4) Why is this algorithm used and not others?

This algorithm uses the Z-coupled channels, which provide the highest signal-to-noise ratio (SNR). This algorithm ensures the highest P_d.

(5) On what principles is the algorithm based (e.g., statistical models, heuristic rules, etc.)?

This algorithm is based on a physical model that exploits the magnetic dipole behavior of EMI anomalies.

(6) What tunable parameters (if any) are used in the detection process (e.g., threshold on signal amplitude, window length, filter coefficients, etc.)?

Detection threshold and radius can be selected.

(7) What are the final values of all tunable parameters for the detection algorithm?

Detection parameters are selected based on the analysis of calibration data. Threshold is set as a function of the noise standard deviation measured in the calibration area. Detection radius is based on the gradients of the smallest target of interest (TOI) near the surface and the largest TOI at depth.

b. <u>Parameter Estimation</u>**.** This section should include the details of which parameters will be extracted from the sensor data for each detected item for characterization. Please answer the following questions:

(1) Which characteristics will be extracted from each detected item and input to the discrimination algorithm (e.g., depth, size, polarizability coefficients, fit quality, etc.)?

Polarizability, size, symmetry, and decay are the parameters required for discrimination.

(2) Why have these characteristics been chosen and not others (e.g., empirical evidence of their ability to help discriminate, inclusion in a theoretical tradition, etc.)?

These features have been proven effective for munitions classification and discrimination at a number of live site demonstrations using similar advanced EMI technologies.

(3) How are these characteristics estimated (e.g., least-mean-squares fit to a dipole model, etc.), include the equations that are used for parameter estimation?

All features are derived from a least-squares fit to a dipole model. The bases for the discrimination features are the object polarizabilities. Polarizabilities are estimated from a linear least-squares inversion of the dipole forward model:

$$\left[H'_x \quad H'_y \quad H'_z\right] = \frac{1}{4\pi R^3}\left[m_x \quad m_y \quad m_z\right]\begin{bmatrix} \frac{3x^2}{R^2}-1 & \frac{3xy}{R^2} & \frac{3xz}{R^2} \\ \frac{3xy}{R^2} & \frac{3y^2}{R^2}-1 & \frac{3yz}{R^2} \\ \frac{3xz}{R^2} & \frac{3yz}{R^2} & \frac{3z^2}{R^2}-1 \end{bmatrix}$$

where m_x, m_y, and m_z are the object principal polarizabilities scaled by the transmitter field:

$$\left[m_x \quad m_y \quad m_z\right] = \left[L_x \quad L_y \quad L_z\right]\begin{bmatrix} H_{Tx}' & 0 & 0 \\ 0 & H_{Ty}' & 0 \\ 0 & 0 & B_{Tz}' \end{bmatrix}$$

The primed coordinates denote the target frame of reference where the magnetic field data are transformed using the Euler rotation angles φ, θ, ψ:

$$\left[H_x' \quad H_y' \quad H_z'\right] = \left[H_x \quad H_y \quad H_z\right]\begin{bmatrix} \cos(\phi)\cos(\theta) & \cos(\phi)\sin(\theta)\sin(\psi)-\sin(\phi)\cos(\psi) & \cos(\phi)\sin(\theta)\cos(\psi)+\sin(\phi)\sin(\psi) \\ \sin(\phi)\cos(\theta) & \sin(\phi)\sin(\theta)\sin(\psi)+\cos(\phi)\cos(\psi) & \sin(\phi)\sin(\theta)\cos(\psi)-\cos(\phi)\sin(\psi) \\ -\sin(\theta) & \cos(\theta)\sin(\psi) & \cos(\theta)\cos(\psi) \end{bmatrix}$$

The rotation angles and the target location (x, y, z) are estimated using a non-linear least squares inversion.

(4) What tunable parameters (if any) are used in the characterization process (e.g., thresholds on background noise, etc.)?

The number of sources and the initial model inputs may be selected.

c. Classification**.** This section should include the details describing the algorithm and associated data and parameters used for discrimination by answering the following questions:

(1) What algorithm is used for discrimination (e.g., multi-layer perception, support vector machine, etc.)?

The primary discrimination method uses a least-squares fit to library polarizabilities. Secondary discrimination methods apply a Gaussian mixture model to the 2-D feature spaces generated from the discrimination parameters.

(2) Why is this algorithm used and not others?

This algorithm has been proven effective for munitions classification and discrimination at a number of live site demonstrations using similar advanced EMI technologies.

(3) Which parameters are considered as possible inputs to the algorithm?

Polarizability, size, symmetry, and decay are the input parameters.

(4) What are the outputs of the algorithm (probabilities, confidence levels)?

The output of the classification algorithm is a list of anomalies ranked in order of confidence level (i.e., highest confidence TOI at top, highest confidence non-TOI at bottom).

(5) How is the threshold set to decide where the munitions/nonmunitions line lies in the discrimination process?

The stop-dig point is selected based on the distance to the feature space cluster centroids and the library fit coefficient.

d. Training. This section should include the details of how training data are used to make a decision on the likelihood of the anomaly correspondence to munitions. Please answer the following questions:

(1) Which tunable parameters have final values that are optimized over a training set of data and which have values that are set according to geophysical knowledge (i.e., intuition, experience, common sense)?

All parameters are set according to a set of training data. Classification decisions are based on variances for model parameters (i.e., polarizabilities, size, decay, symmetry); detection thresholds are based on calibration data noise standard deviation and calibration data TOI type and depth.

(a) For those tunable parameters with final values set according to geophysical knowledge:

1 What is the reasoning behind choosing these particular values?

Not applicable.

2 Why were the final values not optimized over a training set of data?

Not applicable.

(b) For those tunable parameters with final values optimized over the training set data:

1 What training data are used (e.g., all data, a randomly chosen portion of data, etc.)?

All calibration data corresponding to relevant TOIs are used for training.

2 What error metric is minimized during training (e.g., mean squared error, etc.)?

The mean squared error between the forward model output and the training data is minimized.

3 What learning rule is used during training (e.g., gradient descent, etc.)?

Training data are used to compile TOI library features, which can subsequently be mapped to a 2-D feature space. Library features, such as size, decay, and symmetry are derived from the object principal polarizabilities recovered from the least-squares inversion. Once a sufficient sample of training data is acquired, feature variances are established to set the classification decision thresholds.

4 What criterion is used to stop training (e.g., number of iterations exceeds threshold, good generalization over validation set of data, etc.)?

Training is complete once all features are recovered for the site-specific TOIs.

5 Are all tunable parameters optimized at once or in sequence (in sequence = parameters 1 is held constant at some common sense values while parameter 2 is optimized, and then parameter 2 is held constant at its optimized value while parameter 1 is optimized)?

Tunable parameters are optimized using a 2-D feature space analysis. Features are plotted in several feature space combinations (e.g., size versus decay, symmetry versus. decay, etc.). Features that produce the tightest TOI clusters (i.e., minimal variance) are selected for classification analysis.

(2) What are the final values of all tunable parameters for the characterization process?

Classification features are finalized based on the site-specific calibration data.

2.1.4 Data Submission Format

Data were submitted for scoring in accordance with data submission protocols outlined on the USAEC Web site www.uxotestsites.org. These submitted data are not included in this report in order to protect GT information.

2.1.5 Demonstrator Quality Assurance (QA) and Quality Control (QC) (Provided by Demonstrator)

a. Overview of QC. This section is an overview of the complete QC portion of the QA/QC plan. The QC portion is the description of how systems checks are done by the demonstrator to check on items such as tracking, accuracy, drift, and system performance.

QC activities will include basic instrument verification measurements such as noise and static (spike) tests. These instrument verification measurements will be conducted with the OPTEMA system located in a clean (anomaly free) area of the calibration grid.

The noise test will establish a baseline noise floor for the OPTEMA system. For the noise test, the system will be stationed over the cleared location and set to run in a continuous (dynamic) mode where each measurement data block is recorded as a separate data point (i.e., stacking set to 1). Typically, N=30 or more data-points will be collected. To calculate the OPTEMA noise floor, the standard deviation of the N data points of each time-channel for each of the 126 transmitter-receiver combinations is calculated. This process will help to identify any significant deviations in sensor noise on a day-to-day basis. Significant deviations in data channel noise may be indicative of a hardware fault. An example of instrument noise standard deviation is shown in Figure 2. This test will be performed on an initial basis for baseline reference, and subsequently on a daily basis for verification.

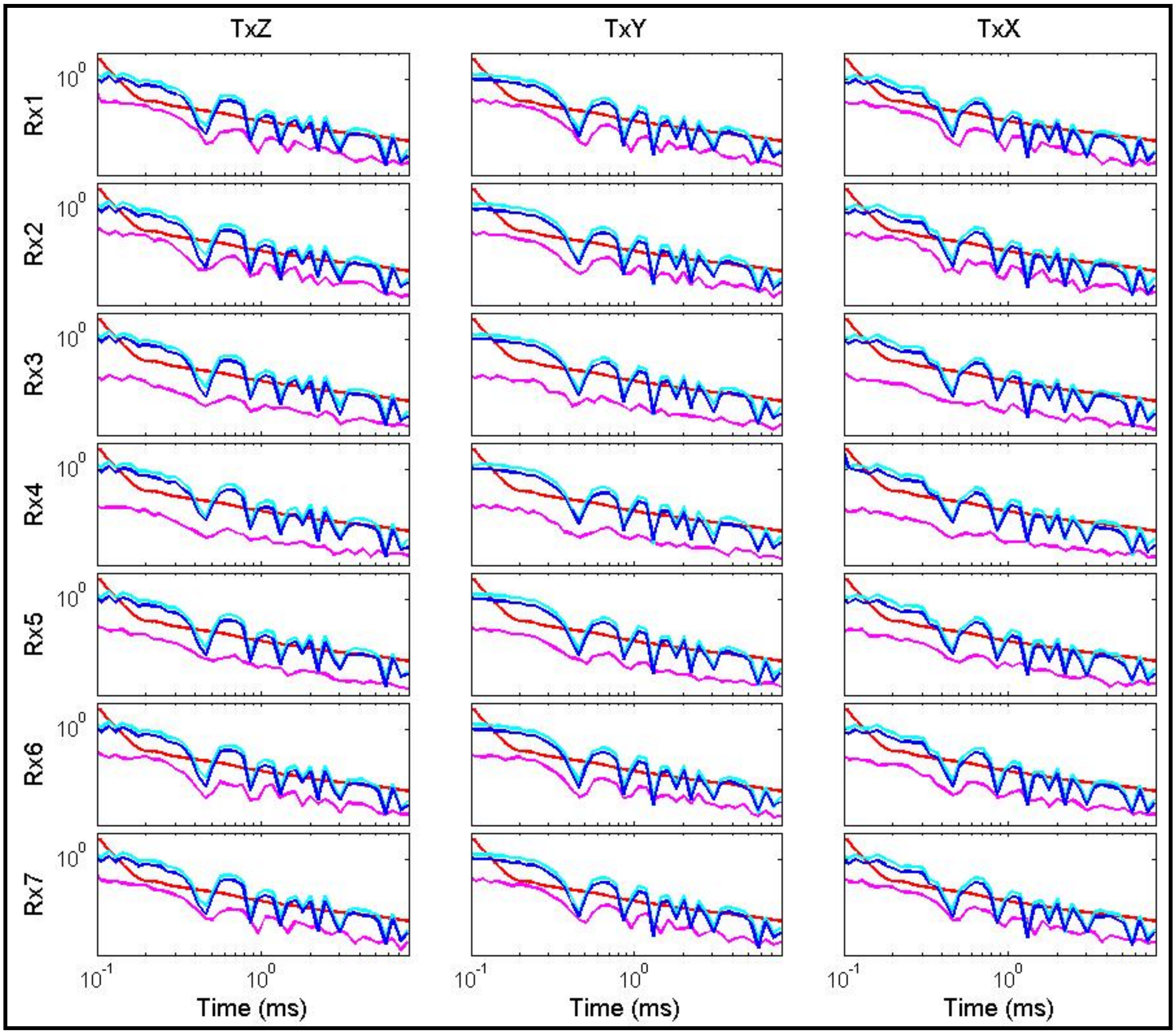


Figure 2. Example of instrument noise standard deviation for seven receiver cubes. The red line represents a baseline quality objective threshold. The magenta, cyan and blue lines correspond to the z-, y-, and x-component receiver data.

Static, or spike, tests will be used to verify consistency in data channel output on a daily basis. Static tests are performed with the OPTEMA stationed in a clean area within the calibration grid. Setting the OPTEMA data acquisition parameters to static mode, a calibration ball is placed in a test jig centered over each receiver cube (fig. 3). This test provides another measurement of instrument consistency. The calibration ball response in each principal data channel (i.e., Transmit-Z/Receive-Z, Transmit-Y/Receive-Y, Transmit-X/Receive-X) should be repeatable within 10-percent deviation on each test (some deviation may be caused by small jig placement inconsistencies). Any significant deviations in the calibration ball response may be indicative of hardware faults. An example of the principal data channel calibration ball response is shown in Figure 4. This test will be performed on an initial basis for baseline reference, and subsequently on a daily basis for verification.

Figure 3. Calibration ball spike test. A calibration ball is placed over each receiver cube in a repeatable location to identify any inconsistencies in data channel output.

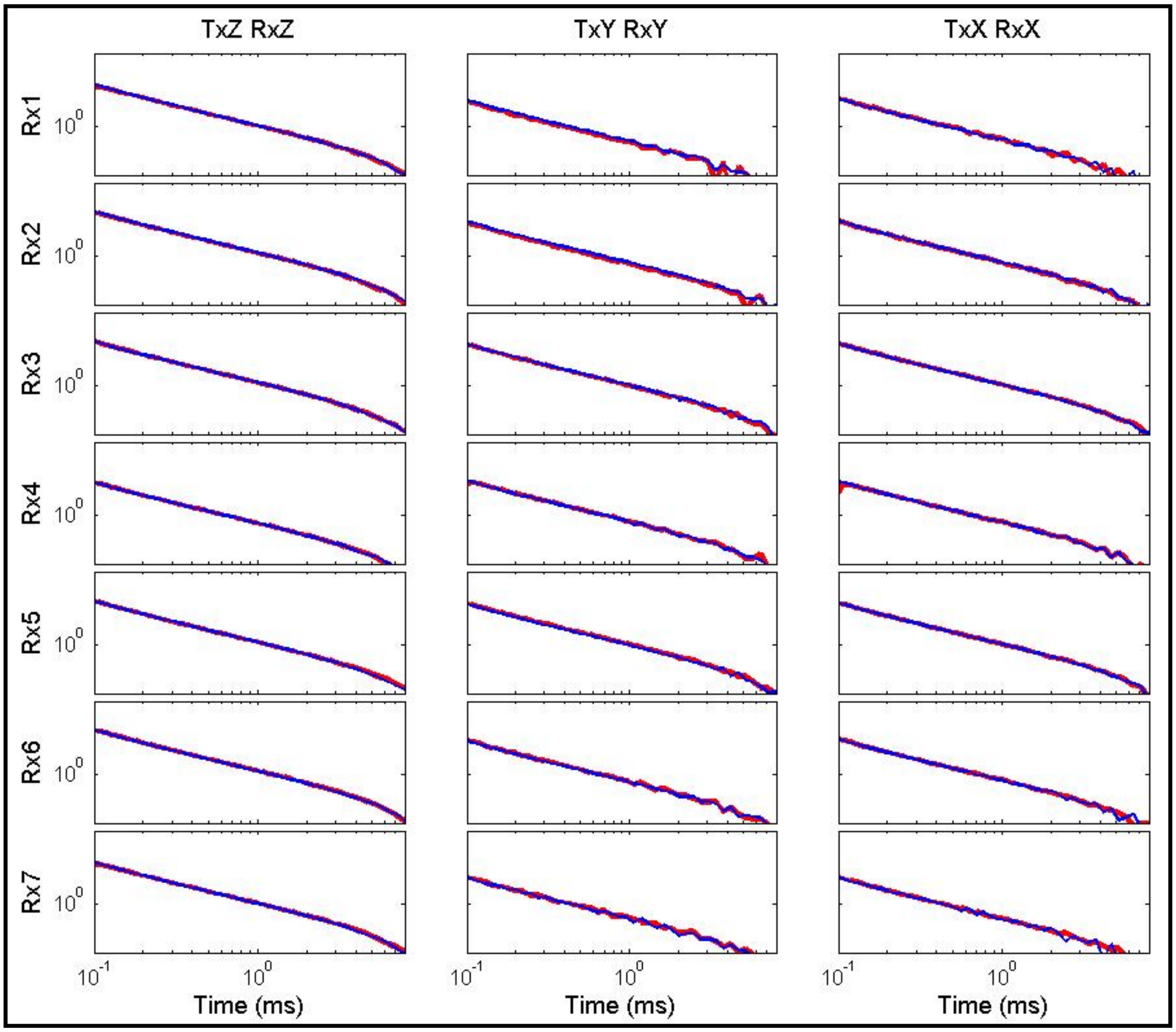


Figure 4. Calibration ball response in principal Tx-Rx pairings for seven receiver cubes. The red line indicates the reference measurement; the blue line indicates the current response.

Initial calibration activities will also include static measurements over calibration grid items. These static data will be used to generate polarizability libraries for TOIs in the blind grid and indirect fire areas. For each TOI type in the calibration grid, we will acquire 3-4 static measurements with each X'-/Y'- pair of transmitters approximately centered over the target. By centering each transmitter pair over the target, we can measure the response across the complete sensor swath; ensuring consistent results are achieved using any subset of data channels within the array.

b. Overview of QA. This section is an overview of the complete QA portion of the QA/QC plan. The QA portion is the description of the procedures to be employed during the demonstration to include items such as lane spacing, sampling rates, and estimated accuracy of navigation and tracking systems.

QA will be ensured through measurement of model output consistency. Static data will be collected over a subset of targets in the calibration grid (three to four targets) regularly to ensure that recovered polarizabilities are consistent on a day-to-day basis. These targets will serve as an instrument verification strip (IVS). Polarizabilities recovered from IVS data should match the reference libraries with 95-percent fit for static measurements collected with each X-/Y-transmitter pair centered over the target. Additionally, estimated target location coordinates recovered from inversion of the IVS data should be within 15 cm of the GT coordinates. Thus, this test serves as a measurement of instrument consistency, model consistency, and positioning accuracy. An example of IVS data matched to library polarizabilities is shown in Figure 5.

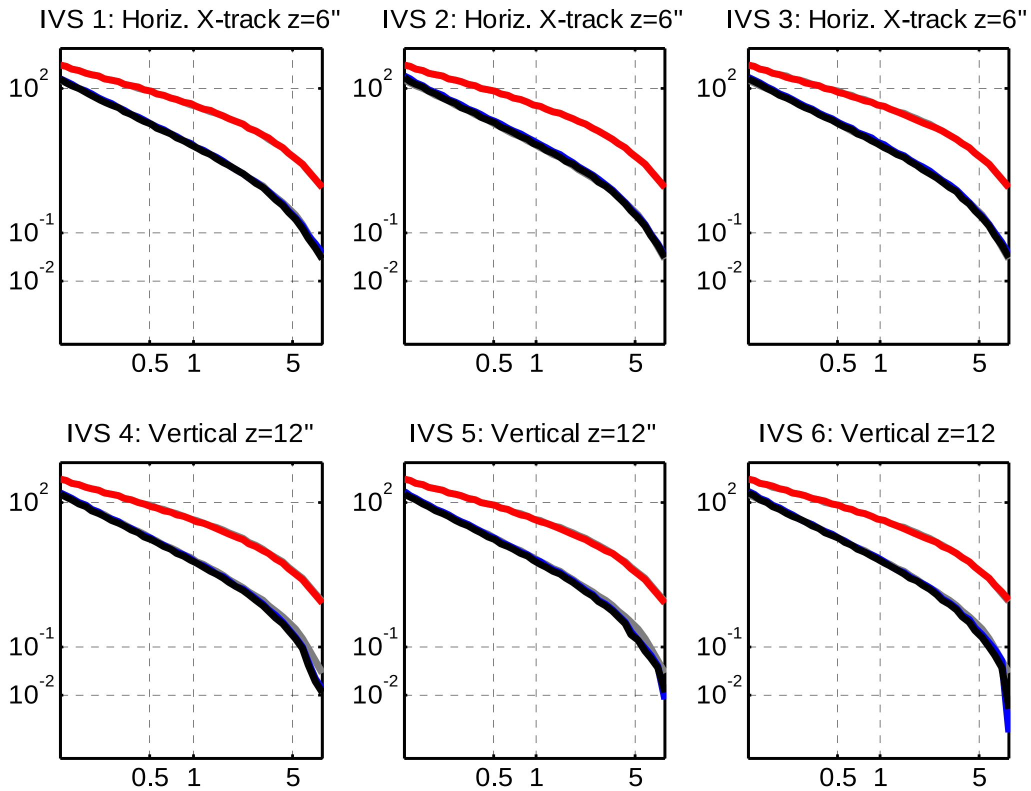


Figure 5. Polarizabilities recovered from daily IVS data matched to a library reference. Red, black and blue lines: primary, secondary and tertiary recovered polarizabilities. Grey lines: reference polarizabilities. IVS items shown here are medium ISOs. All fits are within 95-percent match.

2.1.6 Additional Records

The following record(s) by this vendor can be accessed via the Internet as Microsoft Word documents at www.uxotestsites.org.

2.2 APG SITE INFORMATION

2.2.1 Location

The APG Standardized Test Site is located within a secured range area of the Aberdeen Area. The Aberdeen Area of APG is located approximately 30 miles northeast of Baltimore at the northern end of the Chesapeake Bay. The Standardized Test Site encompasses 17 acres of upland and lowland flats, woods, and wetlands.

2.2.2 Soil Type

According to the soils survey conducted for the entire area of APG in 1998, the test site consists primarily of Elkton Series type soil (ref 2). The Elkton Series consist of very deep, slowly permeable, poorly drained soils. These soils formed in silty aeolin sediments and the underlying loamy alluvial and marine sediments. They are on upland and lowland flats and in depressions of the Mid-Atlantic Coastal Plain. Slopes range from 0 to 2 percent.

The U.S. Army Corps of Engineers Engineering Research and Development Center (ERDC) conducted a site-specific analysis in May 2002 (ref 3). The results basically matched the soil survey mentioned above. Seventy percent of the samples taken were classified as silty loam. The majority (77 percent) of the soil samples had a measured water content between 15 and 30 percent with the water content decreasing slightly with depth.

For more details concerning the soil properties at the APG test site, go to www.uxotestsites.org on the Web to view the entire soils description report.

2.2.3 Test Areas

A description of the test site areas at APG is presented in Table 1. A test site layout is shown in Figure 6.

TABLE 1. TEST SITE AREAS

Area	Description
Calibration lanes	Contains 14 standard munitions items buried in six positions, with representation of clutter, at various angles and depths to allow demonstrators to calibrate their equipment.
Blind grid	Contains 400 grid cells in a 0.5-acre site. The center of each grid cell contains either munitions, clutter, or nothing.
Open field	A 10-acre site composed of generally open and flat terrain with minimal clutter and minor navigational obstacles. Vegetation height varies from 15 to 25 cm. This area is subdivided into four subareas (legacy, direct fire, indirect fire, and challenge).
	• *Open field (legacy)* The legacy subarea contains the same wide variety of randomly-placed munitions that were present in the open field prior to the January 2008 general reconfiguration of the site.
	• *Open field (direct fire)* The direct fire subarea contains only three munition types that could be typically found at an impact area of a direct fire weapons range. Munitions and clutter are placed in a pattern typical for these munitions.
	• *Open field (indirect fire)* The indirect fire subarea contains only three munition types that could be typically found at an impact area of an indirect fire weapons range. Munitions and clutter are placed in a pattern typical for these munitions.
	• *Open field (challenge)* The challenge subarea is easily reconfigurable to meet the specific needs and requirements of the demonstrator or the program sponsor. Any results from this area are not reported in the standardized scoring record.
Woods	1.34-acre area consisting of cleared woods (tree removal with only stumps remaining), partially cleared woods (including all underbrush and fallen trees), and virgin woods (i.e., woods in natural state with all trees, underbrush, and fallen trees left in place).
Moguls	1.30-acre area consisting of two areas (the rectangular or driving portion of the course and the triangular section with more difficult, nondrivable terrain). A series of craters (as deep as 0.91 m) and mounds (as high as 0.91 m) encompass this section.

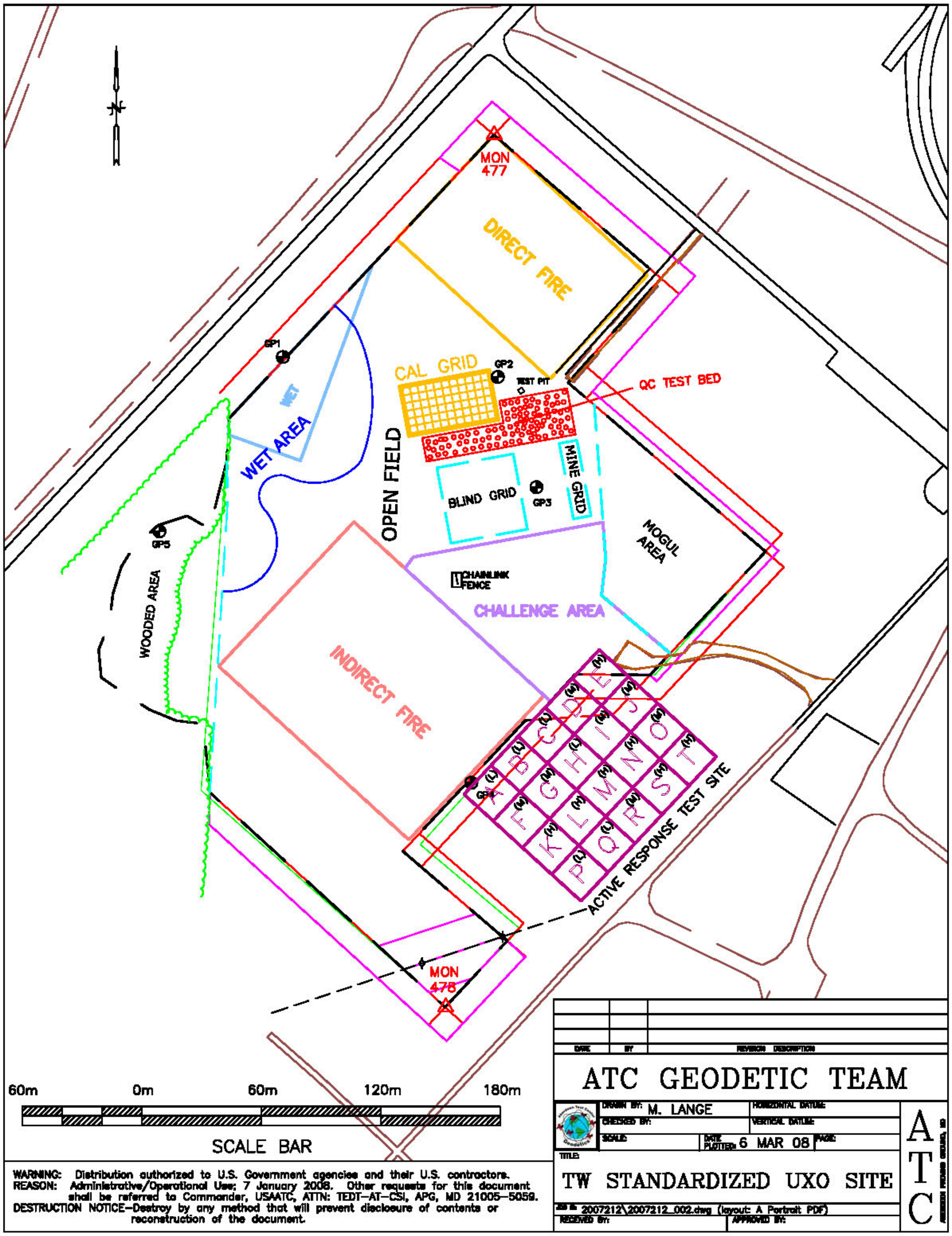


Figure 6. Test site layout.

2.2.4 Standard and Nonstandard Inert Munitions Targets

The standard and nonstandard munitions items emplaced in the test areas are presented in Table 2. Standardized targets are members of a set of specific munitions items that have identical properties to all other items in the set (caliber, configuration, size, weight, aspect ratio, material, filler, magnetic remanence, and nomenclature). Nonstandard targets are inert munitions items having properties that differ from those in the set of standardized items.

TABLE 2. INERT MUNITIONS TARGETS

Item	Munition Type	Calibration Lanes	Blind Grid	Open Field Direct Fire	Open Field Indirect Fire	Open Field Legacy	Moguls	Woods
20-mm projectile M55	S	X				X	X	X
25-mm projectile M794	S	X	X	X				
37-mm projectile M47	S	X	X	X				
40-mm projectile MKII bodies	S	X				X	X	X
BDU-28 submunition	S	X				X	X	X
BLU-26 submunition	S	X				X	X	X
M42 submunition	S	X				X	X	X
57-mm projectile APC M86	S	X				X	X	X
60-mm mortar M49A3	S	X	X		X			
2.75-in. rocket M230	S	X				X	X	X
81-mm mortar M374	S	X	X		X	X	X	X
105-mm HEAT rounds M456	S					X	X	X
105-mm HEAT round M490	S	X	X	X				
105-mm projectile M60	S	X	X		X	X	X	X
155-mm projectile M483A1	S	X				X	X	X
20-mm projectile M55	NS					X	X	X
20-mm projectile M97	NS					X	X	X
40-mm projectile M813	NS					X	X	X
60-mm mortar (JPG)	NS					X	X	X
60-mm mortar M49	NS					X	X	X
2.75-in. rocket M230	NS					X	X	X
2.75-in. rocket XM229	NS					X	X	X
81-mm mortar (JPG)	NS					X	X	X
81-mm mortar M374	NS					X	X	X
105-mm projectile M60	NS					X	X	X
155-mm projectile M483A	NS					X	X	X

APC = Armored personnel carrier.
HEAT = High-explosive antitank.
JPG = Jefferson Proving Ground.
NS = Nonstandard munition.
S = Standard munition.

2.3 ATC SURVEY COMMENTS

None.

SECTION 3. FIELD DATA

3.1 DATE OF FIELD ACTIVITIES (7 THROUGH 11, 14 THROUGH 16 April 2014)

3.2 AREAS TESTED/NUMBER OF HOURS

Areas tested and total numbers of hours operated at each site are presented in Table 3.

TABLE 3. AREAS TESTED AND NUMBER OF HOURS

Area	No. of Hours
Calibration lanes	29.25
Blind grid	5.08
Open field	16.08
Woods	-
Mogul	-
Mine grid	-

Note: Table 3 represents the total time spent in each area.

3.3 TEST CONDITIONS

3.3.1 Weather Conditions

An APG weather station located approximately 1 mile west of the test site was used to record average temperature and precipitation on a half-hour basis for each day of operation. The temperatures presented in Table 4 represent the average temperature during field operations from 0700 to 1700 hours, while precipitation data represent a daily total amount of rainfall. Hourly weather logs used to generate this summary are provided in Appendix B.

TABLE 4. TEMPERATURE/PRECIPITATION DATA SUMMARY

Date, 2014	Average Temperature, °F	Total Daily Precipitation, in.
7 April	45.4	0.40
8 April	59.7	0.06
9 April	57.8	0.00
10 April	58.8	0.00
11 April	68.9	0.01
14 April	71.0	0.00
15 April	61.3	1.07
16 April	39.6	0.00

3.3.2 Field Conditions

White River surveyed the calibration grid, blind grid and a portion of the indirect fire areas. Wet areas from rain prior to and during testing were present. The wet areas made portions of the indirect fire impassable for the OPTEMA system.

3.3.3 Soil Moisture

Three soil probes were placed at various locations within the site to capture soil moisture data: blind grid, calibration, open field, and wooded areas. Measurements were collected in percent moisture and were taken twice daily (morning and afternoon) from five different soil depths (1 to 6 in., 6 to 12 in., 12 to 24 in., 24 to 36 in., and 36 to 48 in.) from each probe. Soil moisture logs are provided in Appendix C.

3.4 FIELD ACTIVITIES

3.4.1 Setup/Mobilization

These activities included initial mobilization and daily equipment preparation and breakdown. A three-person crew took 5 hours, 5 minutes to perform the initial setup and mobilization. A total of 9 hours, 35 minutes of equipment preparation was accrued, and end of day equipment breakdown totaled 3 hours, 30 minutes.

3.4.2 Calibration

White River spent a total of 29 hours, 15 minutes in the calibration lanes, of which 9 hours, 20 minutes were spent collecting data.

3.4.3 Downtime Occasions

Occasions of downtime are grouped into five categories: equipment/data checks or equipment maintenance, equipment failure and repair, weather, demonstration site issues, or breaks/lunch. All downtime is included for the purposes of calculating labor requirements (section 5) except for downtime due to demonstration site issues. Demonstration site issues, while noted in the daily log, are considered nonchargeable downtime for the purposes of calculating labor costs and are not discussed. Breaks and lunches are discussed in this section and billed to the total site survey area.

3.4.3.1 Equipment/data checks, maintenance. Equipment data checks and maintenance activities accounted for 12 hours of site usage time. These activities included changing out batteries and performing routine data checks to ensure the data were being properly recorded/collected. White River spent 25 minutes for breaks and lunches.

3.4.3.2 Equipment failure or repair. No equipment failures occurred during the survey. The vehicle towing the OPTEMA got stuck twice while surveying the indirect fire area, resulting in a brief delay.

3.4.3.3 <u>Weather</u>. Two weather delays occurred during the survey. These weather delays totaled 3 hours, 20 minutes. A badging delay also resulted in 45 minutes of lost survey time.

3.4.4 <u>Data Collection</u>

TABLE 5. TOTAL TIME WHITE RIVER, SPENT PER AREA

Area	Time
Blind grid	3 hr, 10 min
Open field	-
Legacy	-
Direct fire	-
Indirect fire	8 hr, 20 min
Challenge	-
Wooded	-
Mine grid	-
Moguls	-

Note: Table 5 represents the total time spent in each area collecting data.

3.4.5 <u>Demobilization</u>

The White River survey crew conducted a demonstration of the calibration and blind grids. Demobilization occurred on 16 April 2014. On that day, it took the crew 3 hours, 45 minutes to break down and pack up their equipment.

3.5 PROCESSING TIME

White River submitted the raw data from the demonstration activities on the last day of the demonstration, as required. The scoring submittal data were provided 3 July 2014.

3.6 DEMONSTRATOR'S FIELD PERSONNEL

Jon Miller
Erik Russell
Joe Keranen
Greg Schultz

3.7 DEMONSTRATOR'S FIELD SURVEYING METHOD

White River collected the data on a linear basis. The line spacing was the width of the array.

3.8 SUMMARY OF DAILY LOGS

Daily logs capture all field activities during this demonstration and are provided in Appendix D.

SECTION 4. TECHNICAL PERFORMANCE RESULTS

4.1 ROC CURVES USING ALL MUNITIONS CATEGORIES

The probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of clutter detection or probability of false positive within each area are shown in Figures 7 through 12. The probabilities plotted against their respective BAR within each area are shown in Figures 13 through 18. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the GT.

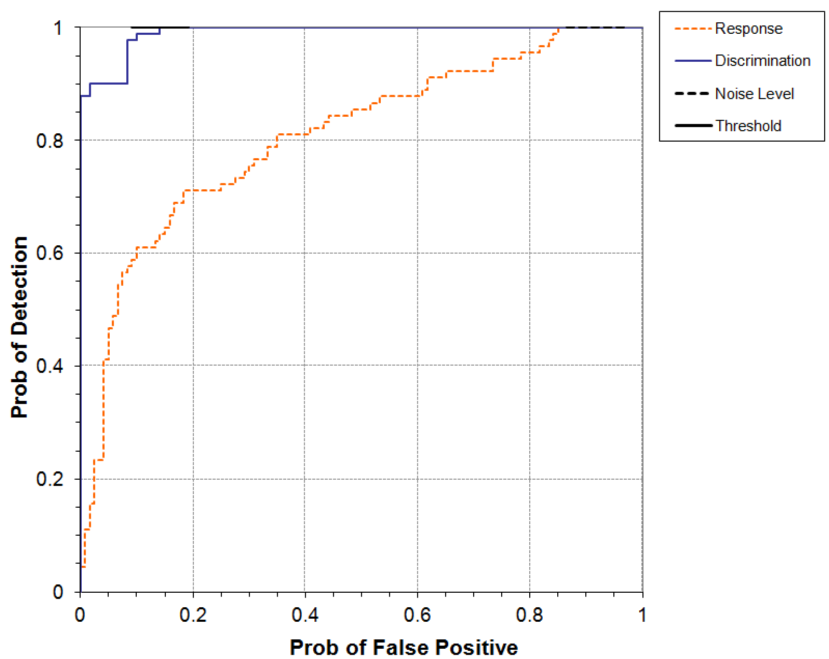


Figure 7. OPTEMA/towed blind grid probability of detection for response and discrimination stages versus their respective probability of false positive.

Not covered

Figure 8. OPTEMA/towed open field (direct fire) probability of detection for response and discrimination stages versus their respective probability of false positive.

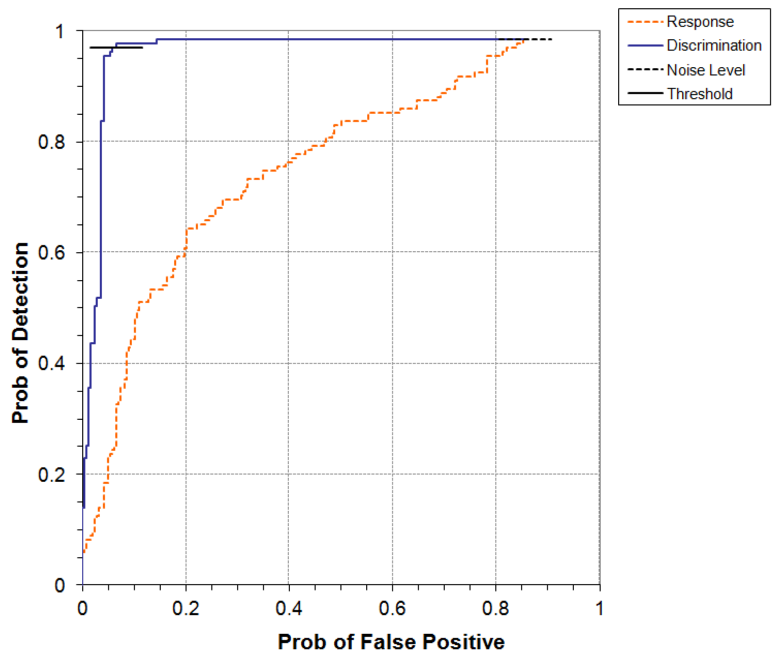


Figure 9. OPTEMA/towed open field (indirect fire) probability of detection for response and discrimination stages versus their respective probability of false positive.

Note: Results for the open field (indirect fire) are for a partial submission.

Not covered

Figure 10. OPTEMA/towed open field (legacy) probability of detection for response and discrimination stages versus their respective probability of false positive.

Not covered

Figure 11. OPTEMA/towed wooded probability of detection for response and discrimination stages versus their respective probability of false positive.

Not covered

Figure 12. OPTEMA/towed mogul probability of detection for response and discrimination stages versus their respective probability of false positive.

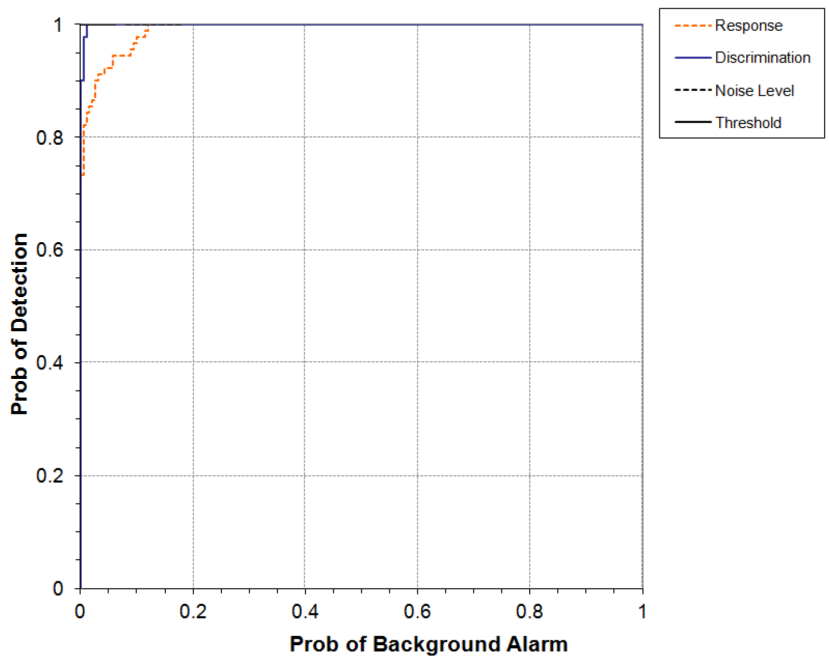
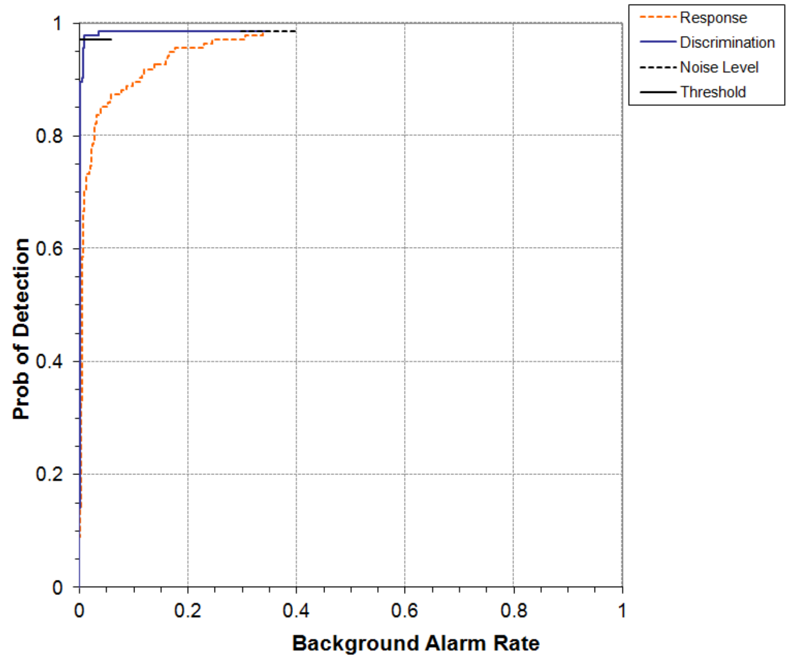


Figure 13. OPTEMA/towed blind grid probability of detection for response and discrimination stages versus their respective probability of background alarm.

Not covered

Figure 14. OPTEMA/towed open field (direct fire) probability of detection for response and discrimination stages versus their respective background alarm rate.



Note: Results for the open field (indirect fire) are for a partial submission.

Figure 15. OPTEMA/towed open field (indirect fire) probability of detection for response and discrimination stages versus their respective background alarm rate.

Not covered

Figure 16. OPTEMA/towed open field (legacy) probability of detection for response and discrimination stages versus their respective background alarm rate.

Not covered

Figure 17. OPTEMA/towed wooded probability of detection for response and discrimination stages versus their respective background alarm rate.

Not covered

Figure 18. OPTEMA/towed mogul probability of detection for response and discrimination stages versus their respective background alarm rate.

4.2 PERFORMANCE SUMMARIES

Results for each of the testing areas are presented in Tables 6a through 6f (labor requirements are provided in section 5). The response stage results are derived from the list of anomalies above the demonstrator-provided noise level. The results for the discrimination stage are derived from the demonstrator's recommended threshold for optimizing munitions related cleanup by minimizing false alarm digs and maximizing munitions recovery. The lower and upper 90-percent confidence limits on P_d, P_{cd}, and P_{fp} were calculated assuming that the number of detections and false positives are binomially distributed random variables.

TABLE 6a. BLIND GRID TEST AREA RESULTS

	Response Stage				**Discrimination Stage**			
Munitions[a] Scores	**Pd^{res}: by type**				**Pd^{disc}: by type**			
	All Types	**105-mm**	**81/60-mm**	**37/25-mm**	**All Types**	**105-mm**	**81/60-mm**	**37/25-mm**
	1.00 1.00 0.98	1.00 1.00 0.93	1.00 1.00 0.93	1.00 1.00 0.93	1.00 1.00 0.98	1.00 1.00 0.93	1.00 1.00 0.93	1.00 1.00 0.93
	By Depth[b]							
0 to 4D	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
4D to 8D	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
8D to 12D	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Clutter Scores	**P_{cd}**				**P_{fp}**			
	By Mass							
By Depth[b]	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**
All Depth	0.95 0.92 0.88	0.83	1.00	1.00	0.19 0.14 0.10	0.02	0.17	0.70
0 to 0.15 m	0.91	0.83	1.00	1.00	0.13	0.02	0.13	1.00
0.15 to 0.3 m	0.94	0.80	1.00	1.00	0.25	0.00	0.43	0.25
0.3 to 0.6 m	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Background Alarm Rates							
	P_{ba}^{res}: 0.13				**P_{ba}^{disc}:** 0.01			

[a]In cells with offset data entries, the numbers to the left are the result and the two numbers to the right are an upper and lower 90-percent confidence interval for an assumed binomial distribution.
[b]All depths are measured to the center of the object.

TABLE 6b. OPEN FIELD DIRECT FIRE TEST AREA RESULTS (NOT COVERED)

	Response Stage				**Discrimination Stage**			
Munitions[a] Scores	***P_d^{res}: by type***				***P_d^{disc}: by type***			
	All Types	**105-mm**	**81-mm**	**60-mm**	**All Types**	**105-mm**	**81-mm**	**60-mm**
	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --
By Density								
High	--	--	--	--	--	--	--	--
Medium	--	--	--	--	--	--	--	--
Low	--	--	--	--	--	--	--	--
By Depth[b]								
0 to 4D	--	--	--	--	--	--	--	--
4D to 8D	--	--	--	--	--	--	--	--
8D to 12D	--	--	--	--	--	--	--	--
Clutter Scores	***P_{cd}***				***P_{fp}***			
By Mass								
By Depth[b]	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**
All Depth	-- -- --	--	--	--	-- -- --	--	--	--
0 to 0.15 m	--	--	--	--	--	--	--	--
0.15 to 0.3 m	--	--	--	--	--	--	--	--
0.3 to 0.6 m	--	--	--	--	--	--	--	--
Background Alarm Rates								
	BAR[res]: --				**BAR[disc]: --**			
Groups								
Found	--				--			
Identified	--				--			
Coverage	--				--			

[a]In cells with offset data entries, the numbers to the left are the result and the two numbers to the right are an upper and lower 90-percent confidence interval for an assumed binomial distribution.
[b]All depths are measured to the center of the object.

TABLE 6c. OPEN FIELD INDIRECT FIRE TEST AREA RESULTS

Response Stage					Discrimination Stage			
Munitions[a] Scores	***P_d^{res}*: by type**				***P_d^{disc}*: by type**			
	All Types	**105-mm**	**81-mm**	**60-mm**	**All Types**	**105-mm**	**81-mm**	**60-mm**
	1.00 0.99 0.96	0.99 0.96 0.88	1.00 1.00 0.95	1.00 1.00 0.95	0.99 0.97 0.94	0.99 0.96 0.88	0.99 0.96 0.89	1.00 1.00 0.95
By Density								
High	0.99	0.96	1.00	1.00	0.99	0.96	1.00	1.00
Medium	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00
Low	0.99	0.95	1.00	1.00	0.96	0.95	0.93	1.00
By Depth[b]								
0 to 4D	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
4D to 8D	1.00	1.00	1.00	1.00	0.96	1.00	0.92	1.00
8D to 12D	0.85	0.33	1.00	1.00	0.85	0.33	1.00	1.00
Clutter Scores	***P_{cd}***				***P_{fp}***			
By Mass								
By Depth[b]	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**
All Depth	0.89 0.86 0.82	0.76	0.98	0.93	0.09 0.07 0.05	0.04	0.05	0.24
0 to 0.15 m	0.87	0.79	0.97	0.93	0.06	0.04	0.05	0.27
0.15 to 0.3 m	0.73	0.42	1.00	0.92	0.10	0.00	0.00	0.25
0.3 to 0.6 m	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00
Background Alarm Rates								
BAR^{res}: 0.35					**BAR^{disc}:** 0.01			
Groups								
Found	1.00				0.84			
Identified	0.28				0.08			
Coverage	0.65				0.45			

[a]In cells with offset data entries, the numbers to the left are the result and the two numbers to the right are an upper and lower 90-percent confidence interval for an assumed binomial distribution.
[b]All depths are measured to the center of the object.

Note: Results for the open field (indirect fire) are for a partial submission.

TABLE 6d. OPEN FIELD LEGACY TEST AREA RESULTS (not covered)

	Response Stage				**Discrimination Stage**			
Munitions[a] Scores	***P_d^{res}: by type***				***P_d^{disc}: by type***			
	All Types	**Small**	**Medium**	**Large**	**All Types**	**Small**	**Medium**	**Large**
	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --
By Depth[b]								
0 to 4D	--	--	--	--	--	--	--	--
4D to 8D	--	--	--	--	--	--	--	--
8D to 12D	--	--	--	--	--	--	--	--
> 12D	--	--	--	--	--	--	--	--

Clutter Scores	***P_{cd}***					***P_{fp}***				
	By Mass									
By Depth[b]	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 10 kg**	**> 10 kg**	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**	**< 10kg**
All Depth	-- -- --	--	--	--	--	-- -- --	--	--	--	--
0 to 0.15 m	--	--	--	--	--	--	--	--	--	--
0.15 to 0.3 m	--	--	--	--	--	--	--	--	--	--
0.3 to 0.6 m	--	--	--	--	--	--	--	--	--	--
> 0.6 m	--	--	--	--	--	--	--	--	--	--

	Background Alarm Rates							
	BAR^{res}:				**BAR^{disc}:**			
	Groups							
Found	--				--			
Identified	--				--			
Coverage	--				--			

[a]The two numbers to the right of the all types munitions result are an upper and lower 90-percent confidence interval for an assumed binomial distribution.
[b]All depths are measured to the center of the object.

TABLE 6e. WOODED TEST AREA RESULTS (not covered)

	Response Stage				Discrimination Stage			
Munitions[a] Scores	***P_d^{res}*: by type**				***P_d^{disc}*: by type**			
	All Types	**Small**	**Medium**	**Large**	**All Types**	**Small**	**Medium**	**Large**
	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --
				By Depth[b]				
0 to 4D	--	--	--	--	--	--	--	--
4D to 8D	--	--	--	--	--	--	--	--
8D to 12D	--	--	--	--	--	--	--	--
> 12D	--	--	--	--	--	--	--	--
Clutter Scores	***P_{cd}***				***P_{fp}***			

					By Mass					
By Depth[b]	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 10 kg**	**> 10 kg**	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**	**< 10kg**
All Depth	-- -- --	--	--	--	--	-- -- --	--	--	--	--
0 to 0.15 m	--	--	--	--	--	--	--	--	--	--
0.15 to 0.3 m	--	--	--	--	--	--	--	--	--	--
0.3 to 0.6 m	--	--	--	--	--	--	--	--	--	--
> 0.6 m	--	--	--	--	--	--	--	--	--	--
					Background Alarm Rates					
	BAR^{res}:					**BAR^{disc}:**				
					Groups					
Found	--					--				
Identified	--					--				
Coverage	--					--				

[a]The two numbers to the right of the all types munitions result are an upper and lower 90-percent confidence interval for an assumed binomial distribution.
[b]All depths are measured to the center of the object.

TABLE 6f. MOGUL TEST AREA RESULTS (not covered)

	Response Stage				Discrimination Stage			
Munitions[a] Scores	***P_d^{res}: by type***				***P_d^{disc}: by type***			
	All Types	**Small**	**Medium**	**Large**	**All Types**	**Small**	**Medium**	**Large**
	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --
	By Depth[b]							
0 to 4D	--	--	--	--	--	--	--	--
4D to 8D	--	--	--	--	--	--	--	--
8D to 12D	--	--	--	--	--	--	--	--
> 12D	--	--	--	--	--	--	--	--

Clutter Scores	***P_{cd}***					***P_{fp}***				
	By Mass									
By Depth[b]	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 10 kg**	**> 10 kg**	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**	**< 10kg**
All Depth	-- -- --	--	--	--	--	-- -- --	--	--	--	--
0 to 0.15 m	--	--	--	--	--	--	--	--	--	--
0.15 to 0.3 m	--	--	--	--	--	--	--	--	--	--
0.3 to 0.6 m	--	--	--	--	--	--	--	--	--	--
> 0.6 m	--	--	--	--	--	--	--	--	--	--

	Background Alarm Rates							
	BAR^{res}:				**BAR^{disc}:**			
	Groups							
Found	--				--			
Identified	--				--			
Coverage	--				--			

[a]The two numbers to the right of the all types munitions result are an upper and lower 90-percent confidence interval for an assumed binomial distribution.
[b]All depths are measured to the center of the object.

4.3 EFFICIENCY, REJECTION RATES, AND TYPE CLASSIFICATION

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are presented in Tables 7a through 7f.

TABLE 7a. BLIND GRID EFFICIENCY AND REJECTION RATES

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	1.00	0.85	0.92
With No Loss of P_d	1.00	0.85	0.92

TABLE 7b. OPEN FIELD (DIRECT) EFFICIENCY AND REJECTION RATES (NOT COVERED)

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	--	--	--
With No Loss of P_d	--	--	--

TABLE 7c. OPEN FIELD (INDIRECT) EFFICIENCY AND REJECTION RATES (NOT COVERED)

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	0.99	0.92	0.97
With No Loss of P_d	1.00	0.83	0.90

Note: Results for the open field (indirect fire) are for a partial submission.

TABLE 7d. OPEN FIELD (LEGACY) EFFICIENCY AND REJECTION RATES (NOT COVERED)

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	--	--	--
With No Loss of P_d	--	--	--

TABLE 7e. WOODED EFFICIENCY AND REJECTION RATES (NOT COVERED)

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	--	--	--
With No Loss of P_d	--	--	--

TABLE 7f. MOGUL EFFICIENCY AND REJECTION RATES (NOT COVERED)

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	--	--	--
With No Loss of P_d	--	--	--

At the demonstrator's recommended setting, the munitions items that were detected and correctly discriminated were further scored on whether their correct type could be identified (tables 8a through 8f). Correct type examples include 20-mm projectile, 105-mm HEAT projectile, and 2.75-inch rocket. A list of the standard type declaration required for each munitions item was provided to demonstrators prior to testing. The standard types for the three example items are 20-mmP, 105H, and 2.75-inch.

TABLE 8a. BLIND GRID CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS

Size	Percentage Correct
25-mm	100
37-mm	100
60-mm	100
81-mm	93
105-mm	73
105-mm artillery	93
Overall	93

TABLE 8b. OPEN FIELD DIRECT FIRE CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS (not covered)

Size	Percentage Correct
25-mm	-
37-mm	-
105-mm	-
Overall	-

TABLE 8c. OPEN FIELD INDIRECT FIRE CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS

Size	Percentage Correct
60-mm	88
81-mm	85
105-mm	96
Overall	90

Note: Results for the open field (indirect fire) are for a partial submission.

TABLE 8d. OPEN FIELD LEGACY CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS (NOT COVERED)

Size	Percentage Correct
Small	-
Medium	-
Large	-
Overall	-

TABLE 8e. WOODED CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS (NOT COVERED)

Size	Percentage Correct
Small	-
Medium	-
Large	-
Overall	-

TABLE 8f. MOGUL CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS (NOT COVERED)

Size	Percentage Correct
Small	-
Medium	-
Large	-
Overall	-

4.4 LOCATION ACCURACY

The mean location error and standard deviations appear in Tables 9a through 9f. These calculations are based on average missed distance for munitions correctly identified during the response stage. Depths are measured from the center of the munitions to the surface. For the blind grid, only depth errors are calculated because (X, Y) positions are known to be the centers of the grid square.

TABLE 9a. BLIND GRID MEAN LOCATION ERROR AND STANDARD DEVIATION

	Mean	Standard Deviation
Northing	NA	NA
Easting	NA	NA
Depth	0.032	0.053

NA = Not available.

TABLE 9b. OPEN FIELD DIRECT FIRE MEAN LOCATION ERROR AND STANDARD DEVIATION (NOT COVERED)

	Mean	Standard Deviation
Northing	-	-
Easting	-	-
Depth	-	-

TABLE 9c. OPEN FIELD INDIRECT FIRE MEAN LOCATION ERROR AND STANDARD DEVIATION

	Mean	Standard Deviation
Northing	0.01	0.07
Easting	-0.02	0.06
Depth	0.01	0.07

Note: Results for the open field (indirect fire) are for a partial submission.

TABLE 9d. OPEN FIELD LEGACY MEAN LOCATION ERROR AND STANDARD DEVIATION (NOT COVERED)

	Mean	Standard Deviation
Northing	-	-
Easting	-	-
Depth	-	-

TABLE 9e. WOODED MEAN LOCATION ERROR AND STANDARD DEVIATION (NOT COVERED)

	Mean	Standard Deviation
Northing	-	-
Easting	-	-
Depth	-	-

TABLE 9f. MOGUL MEAN LOCATION ERROR AND STANDARD DEVIATION (NOT COVERED)

	Mean	Standard Deviation
Northing	-	-
Easting	-	-
Depth	-	-

SECTION 5. APPENDIXES

APPENDIX A. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced munitions item.

Detection: An anomaly location that is within R_{halo} of an emplaced munitions item.

Military Munitions (MM): Specific categories of MM that may pose unique explosive safety risks, including UXO as defined in 10 USC 101(e)(5), discarded military munitions (DMM) as defined in 10 USC 2710(e)(2) and/or munitions constituents (e.g., TNT, RDX) as defined in 10 USC 2710(e)(3) that are present in high enough concentrations to pose an explosive hazard.

Emplaced Munitions: A munitions item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., nonmunitions item) buried by the government at a specified location in the test site.

R_{halo}: A predetermined radius about an emplaced item (clutter or munitions) within which an anomaly identified by the demonstrator as being of interest is considered to be a detection of that item. For the purpose of this program, a circular halo 0.5 meters in radius is placed around the center of the object for all clutter and munitions items.

Small Munitions: Caliber of munitions less than or equal to 40 mm (includes 20-mm projectile, 25-mm projectile, 37-mm projectile, 40-mm projectile, submunitions BLU-26, BLU-63, and M42).

Medium Munitions: Caliber of munitions greater than 40 mm and less than or equal to 81 mm (includes 57-mm projectile, 60-mm mortar, 2.75-inch rocket, and 81-mm mortar).

Large Munitions: Caliber of munitions greater than 81 mm (includes 105-mm HEAT, 105-mm projectile, and 155-mm projectile).

Group: Two or more adjacent GT items with overlapping halos.

GT: Ground truth

Response Stage Noise Level: The level that represents the signal level below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the blind grid test area.

Discrimination Stage Threshold: The demonstrator-selected threshold level that is expected to provide optimum performance of the system by retaining all detectable munitions and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for *n* independent trials with the probability *p* of success and the probability *1-p* of failure being the same for each trial. The number of successes *x* observed in the *n* trials is an estimate of *p* and is considered to be a binomially distributed random variable.

RESPONSE AND DISCRIMINATION STAGE DATA

The scoring of the demonstrator's performance is conducted in two stages: response stage and discrimination stage. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver-operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of clutter detection (P_{cd}) or probability of false positive (P_{fp}). Those that do not correspond to any known item are termed background alarms.

The response stage is a measure of whether the sensor can detect an object of interest. For a channel instrument, this value should be closely related to the amplitude of the signal. The demonstrator must report the response level (threshold) below which target responses are deemed insufficient to warrant further investigation. At this stage, minimal processing may be done. This includes filtering long- and short-scale variations, bias removal, and scaling. This processing should be detailed in the data submission.

For a multichannel instrument, the demonstrator must construct a quantity analogous to amplitude. The demonstrator should consider what combination of channels provides the best test for detecting any object that the sensor can detect. The average amplitude across a set of channels is an example of an acceptable response stage quantity. Other methods may be more appropriate for a given sensor. Again, minimal processing can be done, and the demonstrator should explain how this quantity was constructed in their data submission.

The discrimination stage evaluates the demonstrator's ability to correctly identify munitions as such, and to reject clutter. For the same locations as in the response stage anomaly list, the discrimination stage list contains the output of the algorithms applied in the discrimination-stage processing. This list is prioritized based on the demonstrator's determination that an anomaly location is likely to contain munitions. Thus, higher output values are indicative of higher confidence that a munitions item is present at the specified location. For electronic signal processing, priority ranking is based on algorithm output. For other systems, priority ranking is based on human judgment. The demonstrator also selects the threshold that the demonstrator believes will provide optimum system performance, (i.e., that retains all the detected munitions and rejects the maximum amount of clutter).

Note: The two lists provided by the demonstrator contain identical numbers of potential target locations. They differ only in the priority ranking of the declarations.

GROUP SCORING FACTORS

Based on configuration of the GT at the standardized sites and the defined scoring methodology, there exists munitions groups defined as having overlapping halos. In these cases, the following scoring logic is implemented (fig. A-1 through A-9):

a. Overall site scores (i.e., P_d) will consider only isolated munitions and clutter items.

b. GT items that have overlapping halos (both munitions and clutter) will form a group and groups may form chains.

c. Groups will have a complex halos composed of all the composite halos of all its GT items.

d. Groups will have three scoring factors: groups found groups identified and group coverage. Scores will be based on 1:1 matches of anomalies and GT.

(1) Groups Found (Found). The number of groups that have one or more GT items matched divided by the total number of groups. Demonstrators will be credited with detecting a group if any item within the group is matched to an anomaly in their list.

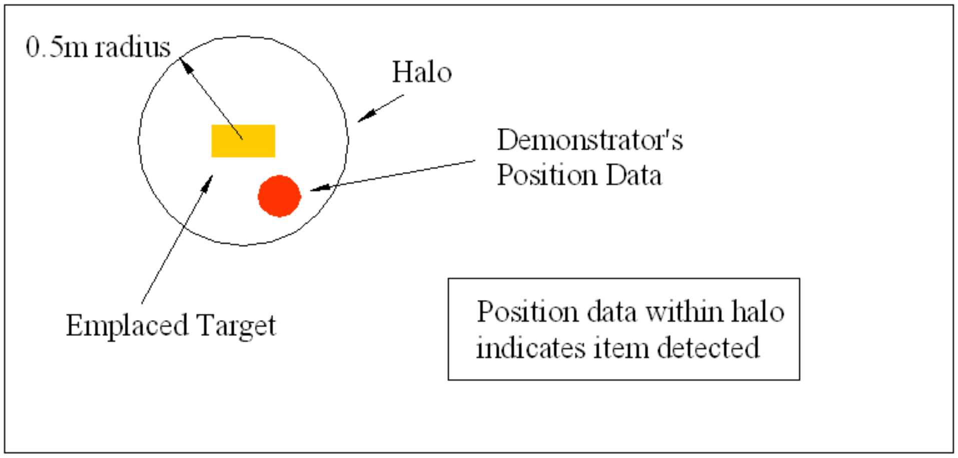
(2) Groups Identified (ID). The number of groups that have two or more GT items matched divided by the total number of groups. Demonstrators will be credited with identifying that a group is present if multiple items within the composite halo are matched to anomalies in their list.

(3) Group Coverage (Coverage). The number of GT items matched within groups divided by the total number of GT items within groups. This metric measures the demonstrator accuracy in determining the number of anomalies within a group. If five items are present and only two anomalies are matched, the demonstrator will score 0.4. If all five are matched the demonstrator will score 1.0.

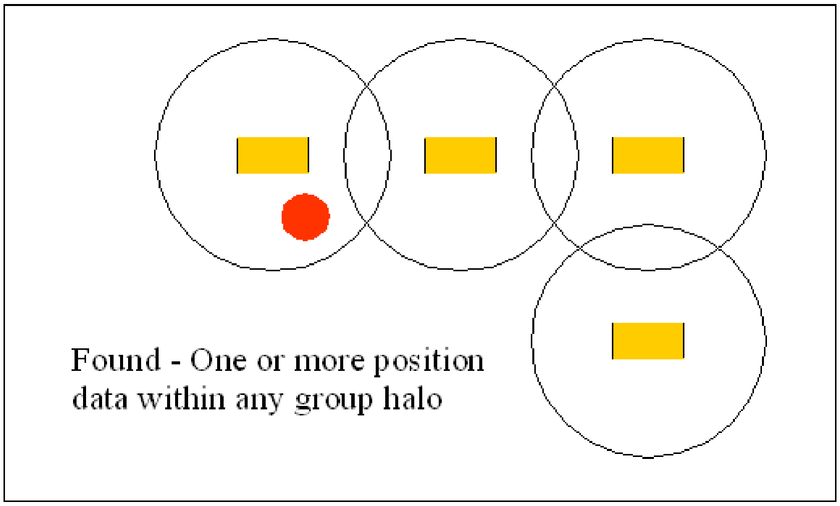
e. Location error will not be reported for groups.

f. Demonstrators will not be asked to call out groups in their scoring submissions. If multiple anomalies are indicated in a small area, the demonstrator will report all individual anomalies.

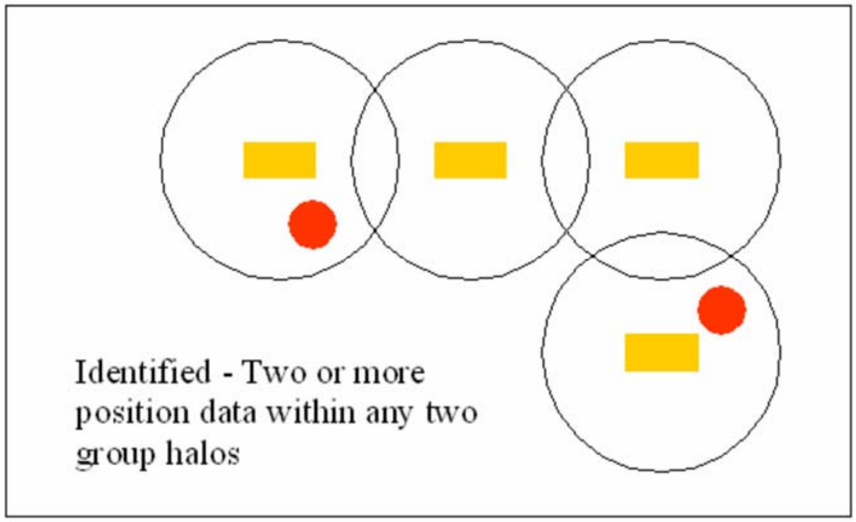
g. Excess alarms within a halo will be disregarded.



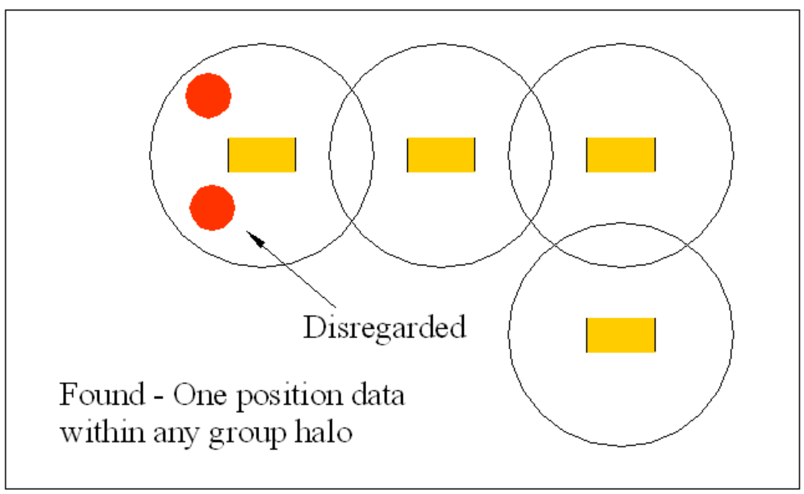
A-1. Example of detected item.



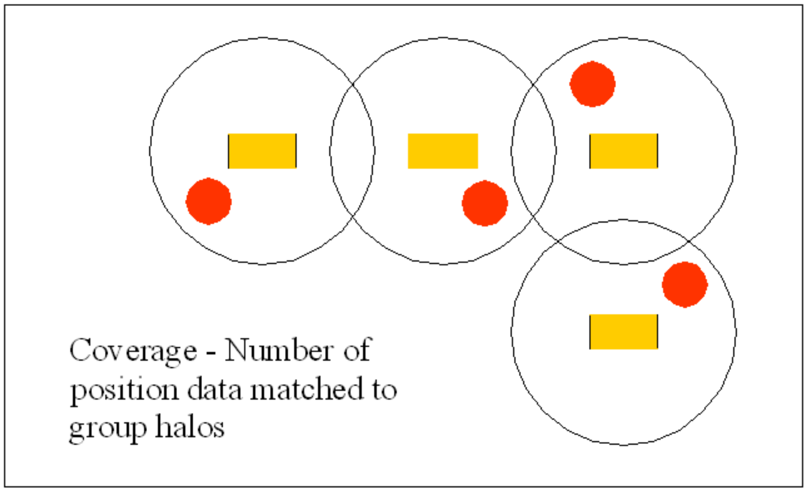
A-2. Example of group found (found).



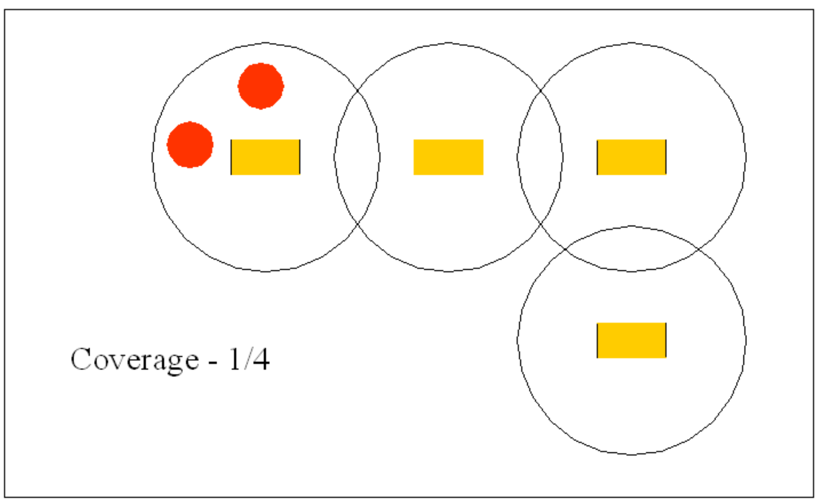
A-3. Example of group identified (ID).



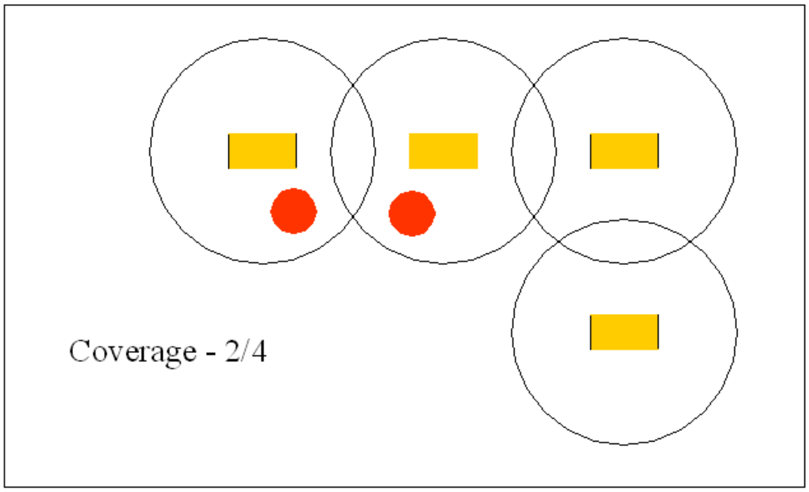
A-4. Example of excess alarms disregarded.



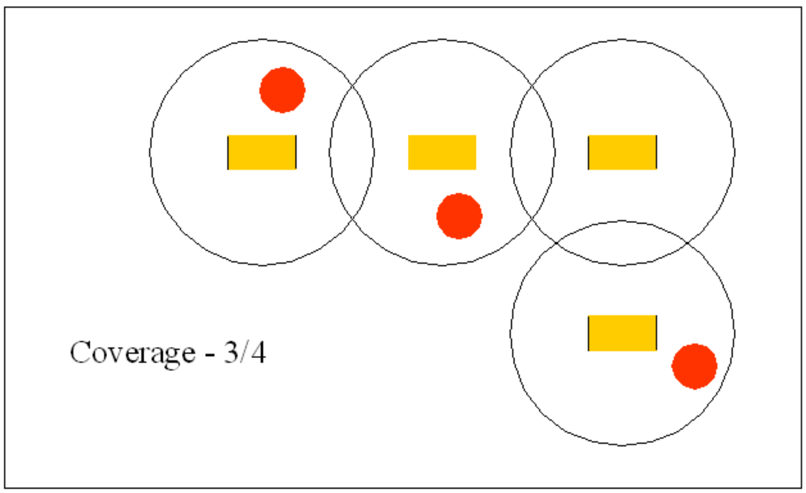
A-5. Example of a group.



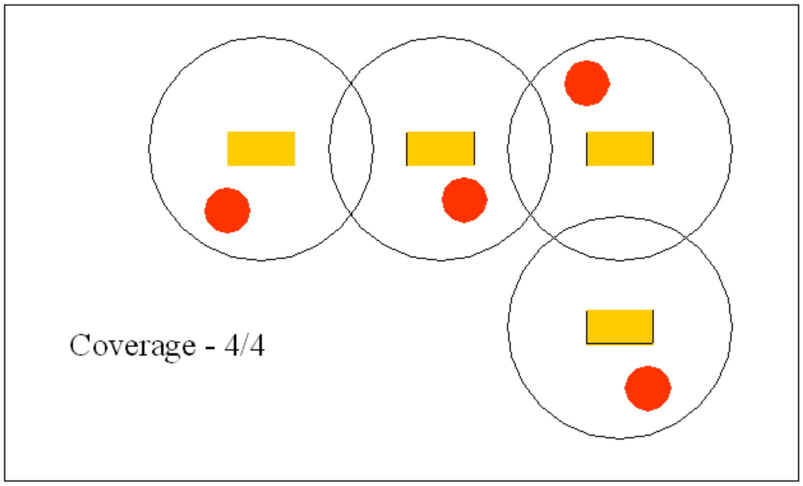
A-6. Example of group (1/4 = 0.25).



A-7. Example of group (2/4 = 0.5).



A-8. Example of group (3/4 = 0.75).



A-9. Example of group (4/4 = 1.0).

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response-stage detections)/(No. of emplaced munitions in the test site).

Response Stage Clutter Detection (cd^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of Clutter Detection (P_{cd}^{res}): P_{cd}^{res} = (No. of response-stage clutter detections)/(No. of emplaced clutter items).

Response Stage Background Alarm (ba^{res}): An anomaly in a blind grid cell that contains neither emplaced munitions nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced munitions or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): Blind grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open field any challenge area (including the direct and indirect firing sub areas) only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{cd}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can therefore be written as $P_d^{res}(t^{res})$, $P_{cd}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to sensor data to discriminate munitions from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to munitions, as well as those that the demonstrator has high confidence correspond to nonmunitions or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination-stage detections)/(No. of emplaced munitions in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm (ba^{disc}): An anomaly in a blind grid cell that contains neither emplaced munitions nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced munitions or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination-stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination-stage signal strength. These quantities can therefore be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER-OPERATING CHARACTERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{cd} or P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] P_d versus P_{fp} and P_d versus BAR being combined into ROC curves are shown in Figure A-10. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

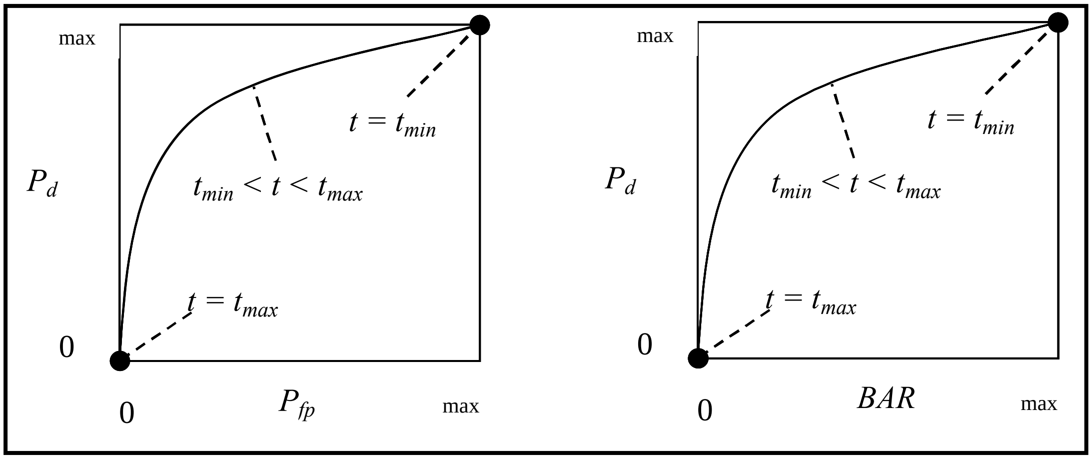


Figure A-10. ROC curves for open field testing. Each curve applies to both the response and discrimination stages.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of munitions detections from the anomaly list while rejecting the maximum number of anomalies arising from nonmunitions items. The efficiency measures the fraction of detected munitions retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum munitions detectable by the sensor and its accompanying clutter detection rate/false positive rate or background alarm rate.

---

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a predetermined and fixed number of detection opportunities (some of the opportunities are located over munitions and others are located over clutter or blank spots). In an open field scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the open field ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the blind grid test sites are true ROC curves.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$: Measures (at a threshold of interest) the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage tmin) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the munitions initially detected in the response stage were retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{cd}^{res}(t_{min}^{res})]$: Measures (at a threshold of interest) the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage tmin). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$.
Open field: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})])$.

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON

The Chi-square test for differences in probabilities (or 2 by 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations.

The test statistic of the 2 by 2 contingency table is the Chi-square distribution with one degree of freedom. When an association between a more challenging terrain feature and relatively degraded performance is sought, a one-sided test is performed. A two-sided 2 by 2 contingency table is used in the Standardized UXO Technology Demonstration Site Program to compare performance between any two areas or subareas when the direction of degradation cannot be predetermined.

For a one-sided test, a significance level of 0.05 is used to set the critical decision limit. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, then the lower proportion tested will be considered significantly less than the greater one (degraded). If the test statistic calculated from the data is less than this value, then no degradation can be said to exist because of the terrain feature introduced.

For a two-sided test, a significance level of 0.10 is used to allow 0.05 on either side of the decision. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, then the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, then the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The Chi-square test cannot be used in these instances. Instead, Fischer's test is used, and the critical decision limit for one-sided tests is the chosen significance level, which in this case is 0.05. With Fischer's test, if the test statistic is less than the critical value, then the proportions are considered to be significantly different.

An example follows that illustrates Standardized UXO Technology Demonstration Site blind grid results compared to those from the open field legacy. It should be noted that a significant result does not prove a cause-and-effect relationship exists between the two populations of interest; however, it does serve as a tool to indicate that one data set has experienced a degradation or change in system performance at a large enough level than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying the blind grid and open field (legacy) using the same system (results indicate the number of munitions detected divided by the number of munitions emplaced):

	Blind grid	Open field
P_d^{res}	100/100 = 1.0	8/10 = .80

P_d^{res}: BLIND GRID versus OPEN FIELD (legacy). Using the example data above to compare probabilities of detection in the response stage, all 100 munitions out of 100 emplaced munitions items were detected in the blind grid while 8 munitions out of 10 emplaced were detected in the open field. Fischer's test must be used since a 100-percent success rate occurs in the data. Fischer's test uses the four input values to calculate a test statistic of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause-and-effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open field relative to results from the blind grid using the same system. This is an example of a one-sided Chi-squared test.

APPENDIX B. DAILY WEATHER LOGS

Date, 2014	Time, EST	Average Temperature, °F	Total Precipitation, in.
7 April	0700	41.2	0.00
	0800	41.5	0.00
	0900	42.2	0.01
	1000	42.3	0.02
	1100	42.7	0.02
	1200	43.8	0.04
	1300	45.0	0.00
	1400	47.3	0.01
	1500	50.1	0.00
	1600	51.2	0.19
	1700	51.8	0.08
8 April	0700	53.9	0.00
	0800	54.2	0.00
	0900	55.9	0.01
	1000	57.6	0.00
	1100	60.0	0.00
	1200	61.9	0.00
	1300	63.2	0.00
	1400	63.1	0.00
	1500	62.9	0.00
	1600	62.5	0.00
	1700	61.5	0.00
9 April	0700	41.0	0.00
	0800	49.8	0.00
	0900	53.5	0.00
	1000	56.4	0.00
	1100	58.9	0.00
	1200	60.8	0.00
	1300	61.9	0.00
	1400	62.4	0.00
	1500	63.4	0.00
	1600	63.6	0.00
	1700	64.2	0.00
10 April	0700	38.0	0.00
	0800	49.4	0.00
	0900	53.6	0.00
	1000	56.5	0.00
	1100	59.9	0.00
	1200	62.3	0.00
	1300	63.6	0.00
	1400	65.3	0.00
	1500	65.7	0.00
	1600	66.2	0.00
	1700	65.9	0.00

EST = Eastern Standard Time.

Date, 2014	Time, EST	Average Temperature, °F	Total Precipitation, in.
11 April	0700	56.7	0.00
	0800	60.0	0.00
	0900	63.6	0.00
	1000	65.8	0.00
	1100	67.6	0.00
	1200	70.6	0.00
	1300	72.9	0.00
	1400	74.9	0.00
	1500	76.6	0.00
	1600	76.9	0.00
	1700	72.0	0.00
14 April	0700	63.6	0.00
	0800	65.5	0.00
	0900	67.9	0.00
	1000	68.4	0.00
	1100	70.9	0.00
	1200	71.9	0.00
	1300	73.8	0.00
	1400	75.0	0.00
	1500	76.1	0.00
	1600	74.0	0.00
	1700	73.6	0.00
15 April	0700	66.8	0.00
	0800	67.3	0.01
	0900	66.7	0.10
	1000	63.2	0.53
	1100	63.2	0.11
	1200	66.1	0.00
	1300	68.1	0.00
	1400	65.6	0.05
	1500	54.2	0.02
	1600	48.0	0.06
	1700	44.9	0.01
16 April	0700	32.1	0.00
	0800	33.2	0.00
	0900	34.7	0.00
	1000	36.3	0.00
	1100	38.1	0.00
	1200	40.1	0.00
	1300	42.2	0.00
	1400	43.5	0.00
	1500	44.5	0.00
	1600	45.5	0.00
	1700	45.6	0.00

[a]Eastern Standard Time.

APPENDIX C. SOIL MOISTURE

Date: 7 April 2014 **Time:** 1600			
Probe Location	**Layer, in.**	**A.M. Reading, %**	**P.M. Reading, %**
Wet area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Wooded area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Open area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Calibration lanes	0 to 6		21.5
	6 to 12		29.9
	12 to 24		32.3
	24 to 36		35.1
	36 to 48		54.6
Blind grid/moguls	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-

Date: 8 April 2014
Time: 0715, 1430

Probe Location	Layer, in.	A.M. Reading, %	P.M. Reading, %
Wet area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Wooded area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Open area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Calibration lanes	0 to 6	21.6	21.6
	6 to 12	29.7	29.6
	12 to 24	32.0	31.9
	24 to 36	35.7	35.5
	36 to 48	54.5	54.7
Blind grid/moguls	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-

Date: 9 April 2014 **Time:** 0800, 1600			
Probe Location	**Layer, in.**	**A.M. Reading, %**	**P.M. Reading, %**
Wet area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Wooded area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Open area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Calibration lanes	0 to 6	21.4	21.3
	6 to 12	29.5	29.4
	12 to 24	31.6	31.6
	24 to 36	35.2	35.3
	36 to 48	54.4	54.3
Blind grid/moguls	0 to 6		
	6 to 12		
	12 to 24		
	24 to 36		
	36 to 48		

Date: 10 April 2014 **Time:** 0800, 1600			
Probe Location	**Layer, in.**	**A.M. Reading, %**	**P.M. Reading, %**
Wet area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Wooded area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Open area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Calibration lanes	0 to 6	21.1	21.0
	6 to 12	29.3	29.2
	12 to 24	31.5	31.5
	24 to 36	35.4	35.2
	36 to 48	54.1	54.0
Blind grid/moguls	0 to 6		
	6 to 12		
	12 to 24		
	24 to 36		
	36 to 48		

Date: 11 April 2014
Time: 1130

Probe Location	Layer, in.	A.M. Reading, %	P.M. Reading, %
Wet area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Wooded area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Open area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Calibration lanes	0 to 6	20.7	
	6 to 12	28.9	
	12 to 24	31.4	
	24 to 36	35.3	
	36 to 48	53.8	
Blind grid/moguls	0 to 6		
	6 to 12	16.8	
	12 to 24	24.9	
	24 to 36	29.5	
	36 to 48	30.2	

Date: 14 April 2014
Time: 0730, 1600

Probe Location	Layer, in.	A.M. Reading, %	P.M. Reading, %
Wet area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Wooded area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Open area	0 to 6	19.8	19.7
	6 to 12	30.6	30.4
	12 to 24	34.8	34.7
	24 to 36	37.9	37.9
	36 to 48	47.6	47.4
Calibration lanes	0 to 6		
	6 to 12	20.4	
	12 to 24	28.5	
	24 to 36	31.1	
	36 to 48	35.2	
Blind grid/moguls	0 to 6	53.6	
	6 to 12		
	12 to 24	16.5	
	24 to 36	24.7	
	36 to 48	29.2	

Date: 15 April 2014
Time: 1600

Probe Location	Layer, in.	A.M. Reading, %	P.M. Reading, %
Wet area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Wooded area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Open area	0 to 6	-	23.6
	6 to 12	-	35.8
	12 to 24	-	38.6
	24 to 36	-	40.2
	36 to 48	-	48.9
Calibration lanes	0 to 6		
	6 to 12		
	12 to 24		
	24 to 36		
	36 to 48		
Blind grid/moguls	0 to 6		
	6 to 12		
	12 to 24		
	24 to 36		
	36 to 48		

Date: 16 April 2014
Time: 0800, 1600

Probe Location	Layer, in.	A.M. Reading, %	P.M. Reading, %
Wet area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Wooded area	0 to 6	-	-
	6 to 12	-	-
	12 to 24	-	-
	24 to 36	-	-
	36 to 48	-	-
Open area	0 to 6	23.8	
	6 to 12	35.9	
	12 to 24	38.5	
	24 to 36	40.6	
	36 to 48	49.7	
Calibration lanes	0 to 6		
	6 to 12		
	12 to 24		
	24 to 36		
	36 to 48		
Blind grid/moguls	0 to 6		21.7
	6 to 12		27.7
	12 to 24		32.5
	24 to 36		34.4
	36 to 48		39.5

APPENDIX D. DAILY ACTIVITY LOGS

Date, 2014	No. of People	Area Tested	Status Start Time, hr	Status Stop Time, hr	Duration min.	Operational Status	Operational Status Comments	Track Method	Pattern	Field Conditions	
7 April	3	Calibration Lanes	0830	1335	305	Initial Setup	Initial Mobilization	GPS	Linear	Rainy	Cool
7 April	3	Calibration Lanes	1335	1510	95	Collecting Data	Collecting Data	GPS	Linear	Rainy	Cool
7 April	3	Calibration Lanes	1510	1535	25	Downtime Due to Equipment Maintenance/Check	Data Check	GPS	Linear	Rainy	Cool
7 April	3	Calibration Lanes	1535	1605	30	Collecting Data	Equipment Breakdown	GPS	Linear	Rainy	Cool
8 April	4	Calibration Lanes	0750	1005	135	Daily Start, Stop	Set Up Equipment	GPS	Linear	Cloudy	Cool
8 April	4	Calibration Lanes	1005	1120	75	Collecting Data	Collecting Data	GPS	Linear	Cloudy	Cool
8 April	4	Calibration Lanes	1120	1255	95	Downtime Due to Equipment Maintenance/Check	Data Check	GPS	Linear	Cloudy	Cool
8 April	4	Calibration Lanes	1255	1440	105	Collecting Data	Collecting Data	GPS	Linear	Cloudy	Cool
8 April	4	Calibration Lanes	1440	1550	70	Downtime Due to Equipment Maintenance/Check	Data Check	GPS	Linear	Cloudy	Cool
8 April	4	Calibration Lanes	1550	1615	25	Daily Start, Stop	Equipment Breakdown	GPS	Linear	Cloudy	Cool
9 April	4	Calibration Lanes	0800	1025	145	Daily Start, Stop	Set Up Equipment	GPS	Linear	Sunny	Cool
9 April	4	Calibration Lanes	1025	1115	50	Collecting Data	Collecting Data	GPS	Linear	Sunny	Cool
9 April	4	Calibration Lanes	1115	1140	25	Downtime Due to Equipment Maintenance/Check	Data Check	GPS	Linear	Sunny	Cool
9 April	4	Calibration Lanes	1140	1220	40	Collecting Data	Collecting Data	GPS	Linear	Sunny	Cool
9 April	4	Calibration Lanes	1220	1440	140	Downtime Due to Equipment Maintenance/Check	Data Check	GPS	Linear	Sunny	Cool
9 April	4	Calibration Lanes	1440	1455	15	Collecting Data	Collecting Data	GPS	Linear	Sunny	Cool

GPS - Global Positioning System.

Date, 2014	No. of People	Area Tested	Status Start Time, hr	Status Stop Time, hr	Duration min.	Operational Status	Operational Status Comments	Track Method	Pattern	Field Conditions	
9 April	4	Calibration Lanes	1455	1550	55	Downtime Due to Equipment Maintenance/Check	Data Check	GPS	Linear	Sunny	Cool
9 April	4	Calibration Lanes	1550	1615	25	Daily Start, Stop	Equipment Breakdown	GPS	Linear	Sunny	Cool
10 April	4	Calibration Lanes	0745	0835	50	Daily Start, Stop	Set Up Equipment	GPS	Linear	Sunny	Cool
10 April	4	Calibration Lanes	0835	1000	85	Collecting Data	Collecting Data	GPS	Linear	Sunny	Cool
10 April	4	Calibration Lanes	1000	1115	75	Downtime Due to Equipment Maintenance/Check	Data Check	GPS	Linear	Sunny	Cool
10 April	4	Calibration Lanes	1115	1200	45	Demonstration Site Issue	Badges for New Individuals Joining Survey	GPS	Linear	Sunny	Cool
10 April	4	Calibration Lanes	1200	1445	165	Daily Start, Stop	Data Check	GPS	Linear	Sunny	Cool
10 April	4	Calibration Lanes	1445	1530	45	Collecting Data	Collecting Data	GPS	Linear	Sunny	Cool
10 April	4	Calibration Lanes	1530	1615	45	Downtime Due to Equipment Maintenance/Check	Equipment Breakdown	GPS	Linear	Sunny	Cool
11 April	4	Calibration Lanes	0740	0825	45	Daily Start, Stop	Set Up Equipment	GPS	Linear	Sunny	Warm
11 April	4	Calibration Lanes	0825	0915	50	Collecting Data	Collecting Data	GPS	Linear	Sunny	Warm
11 April	4	Blind Test Grid	0915	1100	105	Collecting Data	Collecting Data	GPS	Linear	Sunny	Warm
11 April	4	Blind Test Grid	1100	1130	30	Daily Start, Stop	Equipment Breakdown	GPS	Linear	Sunny	Warm
11 April	4	Blind Test Grid	1325	1335	10	Daily Start, Stop	Set Up Equipment	GPS	Linear	Sunny	Warm
11 April	4	Blind Test Grid	1335	1455	80	Collecting Data	Collecting Data	GPS	Linear	Sunny	Warm
14 April	4	Blind Test Grid	1455	1505	10	Downtime Due to Equipment Maintenance/Check	Battery Charge	GPS	Linear	Sunny	Warm
14 April	4	Open Field	1505	1555	50	Collecting Data	Collecting Data/Indirect Fire	GPS	Linear	Sunny	Warm

Date, 2014	No. of People	Area Tested	Status Start Time, hr	Status Stop Time, hr	Duration min.	Operational Status	Operational Status Comments	Track Method	Pattern	Field Conditions	
14 April	4	Open Field	1555	1600	5	Downtime Due to Equipment Maintenance/Check	Charge Batteries	GPS	Linear	Sunny	Warm
14 April	4	Open Field	1600	1615	15	Break/Lunch	Break/Lunch	GPS	Linear	Sunny	Warm
14 April	4	Open Field	0745	0800	15	Collecting Data	Collecting Data/Indirect Fire	GPS	Linear	Sunny	Warm
14 April	4	Open Field	0800	0935	95	Daily Start, Stop	Equipment Breakdown	GPS	Linear	Sunny	Warm
15 April	4	Open Field	0940	1200	140	Weather Issue	Rain, Lightning Advisory	GPS	Linear	Rainy	Cool
15 April	4	Open Field	1200	1350	110	Daily Start, Stop	Set Up Equipment	GPS	Linear	Rainy	Cool
15 April	4	Open Field	1350	1450	60	Weather Issue	Rain, Lightning Advisory	GPS	Linear	Rainy	Cool
15 April	4	Open Field	1450	1515	25	Daily Start, Stop	Equipment Breakdown	GPS	Linear	Rainy	Cool
16 April	4	Open Field	0740	0805	25	Daily Start, Stop	Set Up Equipment	GPS	Linear	Sunny	Cool
16 April	4	Open Field	0805	1145	220	Collecting Data	Collecting Data/Indirect Fire	GPS	Linear	Sunny	Cool
16 April	4	Open Field	1145	1220	35	Downtime Due to Equipment Maintenance/Check	Data Check	GPS	Linear	Sunny	Cool
16 April	4	Blind Test Grid	1220	1230	10	Collecting Data	Collecting Data	GPS	Linear	Sunny	Cool
16 April	4	Blind Test Grid	1230	1615	225	Demobilization	Demobilization	GPS	Linear	Sunny	Cool

APPENDIX E. REFERENCES

1. Standardized UXO Technology Demonstration Site Handbook, DTC Project No. 8-CO-160-000-473, Report No. ATC-8349, March 2002.

2. Aberdeen Proving Ground Soil Survey Report, October 1998.

3. Data Summary, UXO Standardized Test Site: APG Soils Description, May 2002.

APPENDIX F. ABBREVIATIONS

2-D	=	two-dimensional
APC	=	armored personnel carrier
APG	=	Aberdeen Proving Ground
ATC	=	U.S. Army Aberdeen Test Center
ATSS	=	Aberdeen Test Support Services
BAR	=	background alarm rate
DMM	=	discarded military munitions
EMI	=	electromagnetic induction
ERDC	=	U.S. Army Corps of Engineers Engineering Research and Development Center
EST	=	Eastern Standard Time
ESTCP	=	Environmental Security Technology Certification Program
GPS	=	Global Positioning System
GT	=	ground truth
HEAT	=	high-explosive antitank
ID	=	identified
IVS	=	instrument verification strip
JPG	=	Jefferson Proving Ground
MM	=	military munitions
NI	=	National Instruments
NS	=	nonstandard munition
P_{ba}	=	probability of background alarm
P_{cd}	=	probability of clutter detection
P_d	=	probability of detection
P_d^{res}	=	probability of detection for the response stage
P_d^{disc}	=	probability of detection for the discrimination stage
P_{fp}	=	probability of false positive
POC	=	point of contact
QA	=	quality assurance
QC	=	quality control
R_{ba}	=	background alarm rejection
R_{fp}	=	false positive rejection
R_{halo}	=	Halo Radius
ROC	=	receiver-operating characteristic
ROI	=	region of interest
S	=	standard munition
SERDP	=	Strategic Environmental Research and Development Program
SNR	=	signal-to-noise ratio
TDSS	=	Threat Detection and Systems Survivability
TOI	=	target of interest
USAEC	=	U.S. Army Environmental Command
UXO	=	unexploded ordnance
YPG	=	U.S. Army Yuma Proving Ground

APPENDIX G. DISTRIBUTION LIST

ATEC Project No. 2011-DT-ATC-DODSP-F0292

Addressee	No. of Copies
Commander US Army Environmental Command ATTN: IMAE-IT (Dr. Robert Kirgan, K-16) 2450 Connell Road Fort Sam Houston, TX 78234	PDF
Program Manager SERDP/ESTCP Munitions Management ATTN: Mr. Herb Nelson 4800 Mark Center Drive, Suite 17D08 Alexandria, VA 22350	1
White River Technology, Inc. ATTN: Mr. Jonathan Miller 115 Etna Road Building 3, Suite 1 Lebanon, NH 03766	1
Commander U.S. Army Aberdeen Test Center ATTN: TEDT-AT-SLE (Mr. Scott Hill) 400 Colleran Road Aberdeen Proving Ground, MD 21005-5059	1
Defense Technical Information Center 8725 John J. Kingman Road, Suite 0944 Fort Belvoir, VA 22060-6218	PDF

Secondary distribution is controlled by Program Manager, SERDP/ESTCP, Munitions Management, ATTN: Mr. Herb Nelson.